

CC composition for detecting hydrogen peroxide, e.g. for diagnosis of
 CC uraemia, chronic nephritis, giantism and tonic muscular dystrophy. The
 CC enzyme has a lower Km value than prior art creatine amidinohydrolase (cf.
 CC US 5451520). (Updated on 17-OCT-2003 to standardise OS field)

XX Sequence 1212 BP; 244 A; 400 C; 358 G; 210 T; 0 U; 0 Other;

Query Match 99.9%; Score 1210.4; DB 2; Length 1212;
 Best Local Similarity 99.9%; Pred. No. 4.2e-196;
 Matches 1211; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 ATGACTGACGACATGTTGCACTGATGAAATGGCAACAAGCGGAGAAAGTTATTGCGCG 60
DB 1 ATGACTGACGACATGTTGCACTGATGAAATGGCAACAAGCGGAGAAAGTTATTGCGCG 60
QY 61 TTTTGGATGCGAGATGACCGCGCGCAAAAGAGTTGCGGGCTGATGGCCAGAAC 120
DB 61 TTTTGGATGCGAGATGACCGCGCGCAAAAGAGTTGCGGGCTGATGGCCAGAAC 120
QY 121 AATGTGATGCGCGCGCTGTTCACTCTTATCACTGATCACTACTATTCCGGCTGGCTG 180
DB 121 AATGTGATGCGCGCGCTGTTCACTCTTATCACTGATCACTACTATTCCGGCTGGCTG 180
QY 181 TACTGCTATTTGGACGCAAGTAGGCGATGATGACCAACAAGCGCCAGACGATT 240
DB 181 TACTGCTATTTGGACGCAAGTAGGCGATGATGACCAACAAGCGCCAGACGATT 240
QY 241 TGGGCGGCGATGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
DB 241 TGGGCGGCGATGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
QY 301 ACCGACTGCGCGCGCGCAATTTCTATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
DB 301 ACCGACTGCGCGCGCGCGCAATTTCTATCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
QY 361 GCGATGCGCGATGAGTTGCAACAGTCATCTTGACTTCGCGCGCGCGCGCGCGCGCG 420
DB 361 GCGATGCGCGATGAGTTGCAACAGTCATCTTGACTTCGCGCGCGCGCGCGCGCGCG 420
QY 421 CTACCGGGGGTGAAGTTCTGCAATCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
DB 421 CTACCGGGGGTGAAGTTCTGCAATCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
QY 481 TGGCTCGAAGAGAGAGAGTGAATCGCGAAGGCGCGCGCGCGCGCGCGCGCGCGCG 540
DB 481 TGGCTCGAAGAGAGAGAGTGAATCGCGAAGGCGCGCGCGCGCGCGCGCGCGCGCG 540
QY 541 GCGTCGCGCGCTGCGCATCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
DB 541 GCGTCGCGCGCTGCGCATCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
QY 601 AATGCGATGATCGCGAGATCGCAAAATGTTCCCTTCTGTGAGCTGATGACACTGG 660
DB 601 AATGCGATGATCGCGAGATCGCAAAATGTTCCCTTCTGTGAGCTGATGACACTGG 660
QY 661 ACCGTGTTCCATTCGCGCGCATCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
DB 661 ACCGTGTTCCATTCGCGCGCATCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
QY 721 GTGCAATCGCGCGCATCTTTGCTCAACACTTCCGATGATCTTCCGCTACTACAC 780
DB 721 GTGCAATCGCGCGCATCTTTGCTCAACACTTCCGATGATCTTCCGCTACTACAC 780
QY 781 GCGCTGAGAGCGACGCTGTTCTGCGACATGTCGATGATGACGCGCGCGCGCGCGCG 840
DB 781 GCGCTGAGAGCGACGCTGTTCTGCGACATGTCGATGATGACGCGCGCGCGCGCGCG 840
QY 841 AAGAACGTCGCGCTGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
DB 841 AAGAACGTCGCGCTGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
QY 901 GACATGCGCATGACGCTCAACAGATGATCCGCGAGTGGACCTGCTGAAGTACGCTCC 960
DB 901 GACATGCGCATGACGCTCAACAGATGATCCGCGAGTGGACCTGCTGAAGTACGCTCC 960

```

```

DB 901 GACATGCGCATGACGCTCAACAGATGATCCGCGAGTGGACCTGCTGAAGTACGCTCC 960
QY 961 TTGGCTATGACCACTCTTCGCGCGCTGCGCACTACTACTACTGCGAGCGCGCGT 1020
DB 961 TTGGCTATGACCACTCTTCGCGCGCTGCGCACTACTACTACTGCGAGCGCGCGT 1020
QY 1021 GAGCTGCGCGAGACATGACACCGAGCTGAAGCCCGCGAGTGGTCTTCATGAGCCG 1080
DB 1021 GAGCTGCGCGAGACATGACACCGAGCTGAAGCCCGCGAGTGGTCTTCATGAGCCG 1080
QY 1081 ATGCTGATGCTCCCGAGAGGCAATGCCCGGTGCGCGCGCTATGCGAGCAACATCTG 1140
DB 1081 ATGCTGATGCTCCCGAGAGGCAATGCCCGGTGCGCGCGCTATGCGAGCAACATCTG 1140
QY 1141 ATGCTGAGGAGAGAGCGTCCGAGAACATCACCGGCTTCCGTTCCGTCGGAACAAC 1200
DB 1141 ATGCTGAGGAGAGAGCGTCCGAGAACATCACCGGCTTCCGTTCCGTCGGAACAAC 1200
QY 1201 ATCATCCGCAAC 1212
DB 1201 ATCATCCGCAAC 1212

RESULT 3
AAV35699
ID AAV35699 standard; DNA; 1212 BP.
XX
AC AAV35699;
XX
DT 16-SEP-1998 (first entry)
XX
DE Stable creatine amidinohydrazase encoding DNA.
XX
KW Creatine amidinohydrazase; mutant; stable; enzyme; diagnostic agent; ds.
XX
OS Alkaligenes faecalis.
XX
FH Key Location/Qualifiers
FT CDS 1..1212
FT /tag=a
FT /transl_except=(pos:433..435, aa:Glu)
FT /product="Stable creatine amidinohydrazase"
FT /note="the stop codon is not indicated"
XX
PN JP10174585-A.
XX
PD 30-JUN-1998.
XX
PF 17-DEC-1996; 96JP-00337027.
XX
PR 17-DEC-1996; 96JP-00337027.
XX
PA (TOYO) TOYOBO KK.
XX
PA (TOYO) TOYOBO KK.
XX
DR MPI; 1998-421167/36.
XX
DR P-PSDB; AAW61905.
XX
PT New creatine amido-hydrazase used as diagnostic agent - is more stable in
PT neutral buffer than wild type creatine amidino-hydrazase.
XX
PS Disclosure; Page 11-13; 14pp; Japanese.
XX
CC This DNA encodes a stable creatine amidinohydrazase which is a mutant
CC creatine amidinohydrazase and has improved long-term stability in a neutral
CC buffer compared to wild type creatine amidinohydrazase. A recombinant
CC plasmid containing the stable creatine amidinohydrazase gene can be used to
CC transform a cell for the recombinant production of the enzyme. This
CC stable creatine amidinohydrazase is useful as a diagnostic agent can be
CC produced commercially
XX
SQ Sequence 1212 BP; 244 A; 400 C; 358 G; 210 T; 0 U; 0 Other;
Query Match 99.9%; Score 1210.4; DB 2; Length 1212;

```

Best Local Similarity 99.9%; Pred. No. 4,2e-196;
 Matches 1211; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

OY 1 ATGACTGACGACATGTTGACGTGATGAAATGSCACACAGCGGAGAAAGATTATTCGCCG 60
DB 1 ATGACTGACGACATGTTGACGTGATGAAATGSCACACAGCGGAGAAAGATTATTCGCCG 60
OY 61 TTTTGGGATGCGGAGATGACCGCGCGCCAAAAGAGCTTCGGGGCTGGATGGCCAAAGAC 120
DB 61 TTTTGGGATGCGGAGATGACCGCGCGCCAAAAGAGCTTCGGGGCTGGATGGCCAAAGAC 120
OY 121 AATGTCGATGCGGCGCTGTTCACTGTTATCACTGATCAATCACTATTCGGGCTG 180
DB 121 AATGTCGATGCGGCGCTGTTCACTGTTATCACTGATCAATCACTATTCGGGCTG 180
OY 181 TACTGCTATTTGAGCGCAAGTACGAGATGATGACACACACACACACACACGAT 240
DB 181 TACTGCTATTTGAGCGCAAGTACGAGATGATGACACACACACACACACACGAT 240
OY 241 TCGGCGCGGATGACGCGCGCGCCAGCCCTGGCGCGGAGCTTGGGGAACAATCACTAC 300
DB 241 TCGGCGCGGATGACGCGCGCGCCAGCCCTGGCGCGGAGCTTGGGGAACAATCACTAC 300
OY 301 ACCGACTGGCGCGCGGACAAATTTCTATCGGCGCGGAGCTGACACACGCGCGCAAG 360
DB 301 ACCGACTGGCGCGCGGACAAATTTCTATCGGCGCGGAGCTGACACACGCGCGCAAG 360
OY 361 CGCATTCGAGATGAGTTGACACACGTCATCTGACCTTCGCGCGCGAGCTCGAGAGGCC 420
DB 361 CGCATTCGAGATGAGTTGACACACGTCATCTGACCTTCGCGCGCGAGCTCGAGAGGCC 420
OY 421 CTACCGGGGCTGAGTTGCTGACATCACTACGCGCTGATGATGATGATGATGATGAT 480
DB 421 CTACCGGGGCTGAGTTGCTGACATCACTACGCGCTGATGATGATGATGATGATGAT 480
OY 481 TTGGCTCGAAGACAGAGAGTGAATCCGCGAAGGCGCGGCTGATGATGATGATGATGAT 540
DB 481 TTGGCTCGAAGACAGAGAGTGAATCCGCGAAGGCGCGGCTGATGATGATGATGATGAT 540
OY 541 GCGTCGCGGCTGACATCAAGGCGCGGCTGCGCGAGACAGTGAAGTGGCATTCGCCAC 600
DB 541 GCGTCGCGGCTGACATCAAGGCGCGGCTGCGCGAGACAGTGAAGTGGCATTCGCCAC 600
OY 601 AATGCGATGATCGCGAGATCGCAATCGCTTCCTTCGCTGAGAGCTGATGACACTCTG 660
DB 601 AATGCGATGATCGCGAGATCGCAATCGCTTCCTTCGCTGAGAGCTGATGACACTCTG 660
OY 661 AACTGCTTCGATCGCGGACATCAACCGAGCGCGGACAAATCCGCTCAACCGGATC 720
DB 661 AACTGCTTCGATCGCGGACATCAACCGAGCGCGGACAAATCCGCTCAACCGGATC 720
OY 721 GTGCAATCGCGGACATCTCTTCGCTCAACACCTTCGCGATGATCTTCGCGCTACTAC 780
DB 721 GTGCAATCGCGGACATCTCTTCGCTCAACACCTTCGCGATGATCTTCGCGCTACTAC 780
OY 781 GCGCTGAGAGCGACGCTGTTCTGCGACATGTCGATGAGAGCGGACCTTCACACTCGG 840
DB 781 GCGCTGAGAGCGACGCTGTTCTGCGACATGTCGATGAGAGCGGACCTTCACACTCGG 840
OY 841 AAGAACTGCGCGTGTGATCGCGCGGCTCGAGCTATCAAGCGCGGCGCGGCTGCAAG 900
DB 841 AAGAACTGCGCGTGTGATCGCGCGGCTCGAGCTATCAAGCGCGGCGCGGCTGCAAG 900
OY 901 GACATCGCATCGAGCTCAACGAGATGATCCGCGAGTGGAGCTCTGCTGAAGTACCGCTCC 960
DB 901 GACATCGCATCGAGCTCAACGAGATGATCCGCGAGTGGAGCTCTGCTGAAGTACCGCTCC 960
OY 961 TTGGGCTATGCGCATCTCTTCGCGCTGCTGCTGCACTAATCGGTGCGAGGCGCGGCTG 1020
DB 961 TTGGGCTATGCGCATCTCTTCGCGCTGCTGCTGCACTAATCGGTGCGAGGCGCGGCTG 1020
OY 1021 GAGCTGCGGAGGACATCGACACCGAGCTGAAGCGCGGAGTGGTCTTCATGAGCGG 1080
DB 1021 GAGCTGCGGAGGACATCGACACCGAGCTGAAGCGCGGAGTGGTCTTCATGAGCGG 1080

```

```

DB 1021 GAGCTGCGGAGGACATCGACACCGAGCTGAAGCGCGGAGTGGTCTTCATGAGCGG 1080
OY 1081 ATGGTATGCTGCGCGAGGAGCATGCGCGTCCGCGGCTTATCGGACACGACATCTCG 1140
DB 1081 ATGGTATGCTGCGCGAGGAGCATGCGCGTCCGCGGCTTATCGGACACGACATCTCG 1140
OY 1141 ATCGTGGGAGAGACGCTGCGGAGAACATCAACCGGCTTCGCGTTCGCGGACACAC 1200
DB 1141 ATCGTGGGAGAGACGCTGCGGAGAACATCAACCGGCTTCGCGTTCGCGGACACAC 1200
OY 1201 ATCATCGGCAAC 1212
DB 1201 ATCATCGGCAAC 1212

```

RESULT 4
 AAS14742
 ID AAS14742 standard; DNA; 1212 BP.
 XX
 AC AAS14742;
 XX
 DT 19-DEC-2001 (first entry)
 XX
 DS Creatine amidinohydrolase genomic DNA.
 XX
 KM Creatine amidinohydrolase; water; sarcosine; urea; creatinine; uraemia;
 KM chronic nephritis; acute nephritis; tonic muscular dystrophy; giantism;
 KM pigment absorbance; ds.
 XX
 OS Alcaligenes faecalis.
 XX
 FH Key Location/Qualifiers
 FT 1..1212
 FT CDS
 FT /*reg= a
 FT /product= "A. faecalis creatine amidinohydrolase"
 FT /transl_except= (pos:433..435, aa:Glu)
 FT /partial
 FT /note= "No stop codon"
 XX
 PN EP132467-A2.
 XX
 PD 12-SEP-2001.
 XX
 PF 13-FEB-1997; 2001EP-00113052.
 XX
 PR 13-FEB-1996; 96JP-00025435.
 PR 13-FEB-1997; 97EP-00102270.
 XX
 PA (TOYM) TOYO BOSEKI KK.
 XX
 PI Sogabe A, Hattori T, Nishiya Y, Kawamura Y;
 DR WPI; 2001-612481/71.
 DR P-PsDB; AAU08727.
 XX
 PT New creatine amidinohydrolase, useful as a routine reagent for clinical
 PT tests for determining creatine and creatinine in biological samples,
 PT particularly useful in diagnosing diseases such as uremia or chronic
 PT nephritis.
 XX
 PS Disclosure; Page 15-16; 21pp; English.
 XX
 CC The invention relates to Alcaligenes faecalis creatine amidinohydrolase,
 CC which catalyses creatine and water to sarcosine and urea. Creatine
 CC amidinohydrolase can be produced by culturing a microorganism producing
 CC the protein in a nutrient medium and recovering the protein from the
 CC resulting culture. Creatine amidinohydrolase is useful as a routine
 CC reagent for clinical tests for determining creatine and creatinine in
 CC biological samples. This is particularly useful in diagnosing diseases
 CC such as uremia, chronic nephritis, acute nephritis, giantism and tonic
 CC muscular dystrophy. The presence of creatine in a sample can be
 CC determined by measuring an absorbance of a pigment produced by the
 CC reaction of a reagent containing creatine amidinohydrolase with the

CC sample. This sequence represents genomic DNA encoding Alcaligenes
 CC faecalis creatine amidinohydrolase
 XX

Sequence 1212 BP; 244 A; 400 C; 358 G; 210 T; 0 U; 0 Other;

Query Match 99.9%; Score 1210.4; DB 4; Length 1212;

Best Local Similarity 99.9%; Pred. No. 4.2e-196;

Matches 1211; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 ATGACTGACGACATGTTGCACTGATGAATGCGACAAAGCGCGAAGAAATTATTCGCCG 60
DB 1 ATGACTGACGACATGTTGCACTGATGAATGCGACAAAGCGCGAAGAAATTATTCGCCG 60
QY 61 TTTTCGGAATGCGAATGACCGCGCGCAAAAGAGCGTTCGCGGTGATGCGCAAGAAC 120
DB 61 TTTTCGGAATGCGAATGACCGCGCGCAAAAGAGCGTTCGCGGTGATGCGCAAGAAC 120
QY 121 AATGTCGATGCGGCGCTGTTCACTCTTATCACTGATCACTAATTAATTCGCGTGGCTG 180
DB 121 AATGTCGATGCGGCGCTGTTCACTCTTATCACTGATCACTAATTAATTCGCGTGGCTG 180
QY 181 TACTGCTATTTCCGAGCGCAAGTACGGATGTCATGCAACAACGCAAGAGATT 240
DB 181 TACTGCTATTTCCGAGCGCAAGTACGGATGTCATGCAACAACGCAAGAGATT 240
QY 241 TCGCGCGGCGATCGACCGCGCGCGACGCGTGGCGCGGACCTTGGCGGACATCACCTAC 300
DB 241 TCGCGCGGCGATCGACCGCGCGCGACGCGTGGCGCGGACCTTGGCGGACATCACCTAC 300
QY 301 ACCGACTGGCGCGCGCAATTTCTATCGCGCGGTGCGCGAGCTGACACGCGCGCGAAG 360
DB 301 ACCGACTGGCGCGCGCAATTTCTATCGCGCGGTGCGCGAGCTGACACGCGCGCGAAG 360
QY 361 CGGATGCGGATCGAATTTCCACACGTCATCTGATCTTCGCGCGCGAGCTCGAGAAAGCC 420
DB 361 CGGATGCGGATCGAATTTCCACACGTCATCTGATCTTCGCGCGCGAGCTCGAGAAAGCC 420
QY 421 CTACCGGCGGCTGCGATTCGTCGACATCAGCAGCGCTGATGATGATGCGGACATCAAG 480
DB 421 CTACCGGCGGCTGCGATTCGTCGACATCAGCAGCGCTGATGATGATGCGGACATCAAG 480
QY 481 TCGCTCGAAGAGAGAAAGTGCATCGCGAAGCGCGCGCTGTGTGAAGTTCGCGCGCGG 540
DB 481 TCGCTCGAAGAGAGAAAGTGCATCGCGAAGCGCGCGCTGTGTGAAGTTCGCGCGCGG 540
QY 541 GCGTGGCGGCTGCGATCAAGCGCGCGGCTGCGCGAGATGATGATGCGCGCGCGCG 600
DB 541 GCGTGGCGGCTGCGATCAAGCGCGCGGCTGCGCGAGATGATGATGCGCGCGCGCG 600
QY 601 AATGCGATGATCGCGGAGATCGCGAATGATGATGATGATGATGATGATGATGATGATG 660
DB 601 AATGCGATGATCGCGGAGATCGCGAATGATGATGATGATGATGATGATGATGATGATG 660
QY 661 ACTGCTTCAGTTCGCGGATCAACACGCGCGCGCGCAATTCGCGTCAACACCGCATC 720
DB 661 ACTGCTTCAGTTCGCGGATCAACACGCGCGCGCGCAATTCGCGTCAACACCGCATC 720
QY 721 GTGCAATTCGCGGAGATCTTTGCTCAACACTTCCCGATGATCTTCGCGTCACTACCC 780
DB 721 GTGCAATTCGCGGAGATCTTTGCTCAACACTTCCCGATGATCTTCGCGTCACTACCC 780
QY 781 GGGCTGGAAGCGACGCTGTTTTCGCGACATGTCGATGAGAGCGCGCTGCACTATGAGAG 840
DB 781 GGGCTGGAAGCGACGCTGTTTTCGCGACATGTCGATGAGAGCGCGCTGCACTATGAGAG 840
QY 841 AAGAAAGTGGCGCGTGCATCGCGCGGCTGCACTGATCAAGCGCGCGCGCGCTGCAAG 900
DB 841 AAGAAAGTGGCGCGTGCATCGCGCGGCTGCACTGATCAAGCGCGCGCGCGCTGCAAG 900
QY 901 GACATGCGCATGAGCTCAACGAGATGATCGCGGAGTGGGACTGCTGAAGTACCGCTCC 960
DB 901 GACATGCGCATGAGCTCAACGAGATGATCGCGGAGTGGGACTGCTGAAGTACCGCTCC 960

```

```

QY 961 TTGCGTATGCGCACTCTTCGCGCTGCTGTCGCACTACTACGTCGCGAGCGCGCTG 1020
DB 961 TTGCGTATGCGCACTCTTCGCGCTGCTGTCGCACTACTACGTCGCGAGCGCGCTG 1020
QY 1021 GAGCTCGCGAGGACATTCGACACCGAGCTGAAGCCCGCGCATGCTTCTCATGAGCCG 1080
DB 1021 GAGCTCGCGAGGACATTCGACACCGAGCTGAAGCCCGCGCATGCTTCTCATGAGCCG 1080
QY 1081 ATGCTGATGCTCGCGAGGCGCATGCGCGGTCGCGGCGCTATGCGGAGACGACATCCTG 1140
DB 1081 ATGCTGATGCTCGCGAGGCGCATGCGCGGTCGCGGCGCTATGCGGAGACGACATCCTG 1140
QY 1141 ATGCTGCGGAGGACGCTGCCGAGACATCACCGGCTTCCGCTCCGAGCAAC 1200
DB 1141 ATGCTGCGGAGGACGCTGCCGAGACATCACCGGCTTCCGCTCCGAGCAAC 1200
QY 1201 ATCATTCGCAAC 1212
DB 1201 ATCATTCGCAAC 1212

```

RESULT 5

AA199856 standard; DNA; 1212 BP.

AA199856; 07-AUG-2003 (revised)
 28-JAN-2002 (first entry)

Alcaligenes faecalis creatineamidinohydrolase encoding DNA.

Alcaligenes faecalis; TE3581; FERM P14237; creatineamidinohydrolase; ds.

Alcaligenes faecalis.

Key Location/Qualifiers

FT CDS 1..1212 /tag= a

FT FT /transl_except= (pos:433..435,aa:Glu)

FT FT /product= "creatineamidinohydrolase"

FT FT /note= "CDS lacks a stop codon"

PN JP2001252088-A.

PD 18-SEP-2001.

PF 16-MAY-1995; 2001JP-00051054.

PR 16-MAY-1995; 95JP-00117283.

PA (TOYO) TOYOBO KK.

DR WPI; 2002-003140/01.

DR P-PSDB; AAM51471.

XX A gene encoding creatineamidinohydrolase.

PS Claim 4; Page 10; 11pp; Japanese.

CC The invention relates to Alcaligenes faecalis TE3581 (FERM P-14237)

CC creatineamidinohydrolase and the encoding gene. The gene can be used for

CC the commercial preparation of creatineamidinohydrolase. (updated on 07-

CC AUG-2003 to correct OS field.)

XX Sequence 1212 BP; 244 A; 400 C; 358 G; 210 T; 0 U; 0 Other;

XX Query Match 99.9%; Score 1210.4; DB 6; Length 1212;

XX Best Local Similarity 99.9%; Pred. No. 4.2e-196; Mismatches 1; Indels 0; Gaps 0;

XX Matches 1211; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGACTGACGACATGTTGCACTGATGAATGCGACAAAGCGCGAAGAAATTATTCGCCG 60

```

Db      1 ATGACATGACGATGTTGACGTTGATGAAATGCAACAACGCGAGAAATATTATTCGCGG 60
Qy      61 TTTTGGATGCGGAGATGACCCGCGCCCAAAAGAGAGTTCCGGGCTGGATGGCCAAAGAC 120
Db      61 TTTTGGATGCGGAGATGACCCGCGCCCAAAAGAGAGTTCCGGGCTGGATGGCCAAAGAC 120
Qy      121 AATGTCGATGCGGCGCTGTTCACTTCTTATCATCTGATCACTAATTCGGGCTGGCTG 180
Db      121 AATGTCGATGCGGCGCTGTTCACTTCTTATCATCTGATCACTAATTCGGGCTGGCTG 180
Qy      181 TACTGCTATTTTCGAGCGCAAGTACGCGCATGATCATGACCAACAACGCGCAGAGATT 240
Db      181 TACTGCTATTTTCGAGCGCAAGTACGCGCATGATCATGACCAACAACGCGCAGAGATT 240
Qy      241 TCGGCGCGGATGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
Db      241 TCGGCGCGGATGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
Qy      301 ACCGACTGCGCGCGCGCGCAATTCTATGCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
Db      301 ACCGACTGCGCGCGCGCGCAATTCTATGCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
Qy      361 CGCATCGGATGAGTTTCGACCATGACCGCATCTCGACTTCGCGCGCGCGCGCGCGCG 420
Db      361 CGCATCGGATGAGTTTCGACCATGACCGCATCTCGACTTCGCGCGCGCGCGCGCGCG 420
Qy      421 CTACCGGCGGCTGAGTTTCGACCATGACCGCATCTCGACTTCGCGCGCGCGCGCGCG 480
Db      421 CTACCGGCGGCTGAGTTTCGACCATGACCGCATCTCGACTTCGCGCGCGCGCGCGCG 480
Qy      481 TCGCTCGAAGACAGAAAGCTGATTCGCGGAGGCGCGCGCGCGCGCGCGCGCGCGCG 540
Db      481 TCGCTCGAAGACAGAAAGCTGATTCGCGGAGGCGCGCGCGCGCGCGCGCGCGCGCG 540
Qy      541 GCGTCGCGGCTGCGCATCAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
Db      541 GCGTCGCGGCTGCGCATCAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
Qy      601 AATGCGATGATCCGCGAGATCCGCAATGCTTCCCTTGTGAGAGCTGATGACACCTGG 660
Db      601 AATGCGATGATCCGCGAGATCCGCAATGCTTCCCTTGTGAGAGCTGATGACACCTGG 660
Qy      661 ACCTGCTTCATCGCGGATCAACACCGAGGCGCGCAATCCGGTTCACCAACGCGCATC 720
Db      661 ACCTGCTTCATCGCGGATCAACACCGAGGCGCGCAATCCGGTTCACCAACGCGCATC 720
Qy      721 GTGCAATCCGCGGACATCTTTTCGCTCAACACCTTCCGATGATCTTCCGCTACTACAC 780
Db      721 GTGCAATCCGCGGACATCTTTTCGCTCAACACCTTCCGATGATCTTCCGCTACTACAC 780
Qy      781 GCGCTGAGACGACGCTGTTTCGCGACATGTCGATGACGCCAGCTTCGACATCTGGAG 840
Db      781 GCGCTGAGACGACGCTGTTTCGCGACATGTCGATGACGCCAGCTTCGACATCTGGAG 840
Qy      841 AAGAACTGCGCGGATGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
Db      841 AAGAACTGCGCGGATGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
Qy      901 GACATCGCATGAGCTCAACGAGATGATCCGCGAGTGGACCTGCTGAGTACCGCTCC 960
Db      901 GACATCGCATGAGCTCAACGAGATGATCCGCGAGTGGACCTGCTGAGTACCGCTCC 960
Qy      961 TTGCGCTATGCGCATCTCTTCCGCGCTGCTGTCACCTAATCGGTGCGAGCGCGCG 1020
Db      961 TTGCGCTATGCGCATCTCTTCCGCGCTGCTGTCACCTAATCGGTGCGAGCGCGCG 1020
Qy      1021 GAGCTGCGGAGAGACATGACACCGAGCTGAAGCGCGGATGGTCTCCATGAGCGCG 1080
Db      1021 GAGCTGCGGAGAGACATGACACCGAGCTGAAGCGCGGATGGTCTCCATGAGCGCG 1080
Qy      1081 ATGCTGATGCTGCGCGAGGCGCATGCCGCGTGCAGCGGCTATTCGACGACGACATCT 1140

```

```

Db      1081 ATGCTGATGCTGCGGAGGCGCATGCCGCTCCGCGGCTATTCGACACGACATCTCTG 1140
Qy      1141 ATGCTGCGGAGAGACGAGTGCAGACATCAACCGGCTTCGCTTCGCTTCGAGACACAAC 1200
Db      1141 ATGCTGCGGAGAGACGAGTGCAGACATCAACCGGCTTCGCTTCGCTTCGAGACACAAC 1200
Qy      1201 ATCATCCGCAAC 1212
Db      1201 ATCATCCGCAAC 1212

RESULT 6
ABA93696
ID ABA93696 standard; DNA; 1212 BP.
XX
AC ABA93696;
XX
DT 29-APR-2002 (first entry)
XX
DE Creatine amidinohydrazide encoding DNA SEQ ID NO:2.
XX
KM Creatine amidinohydrazide; enzyme; mutant; stable; clinical diagnosis;
XX gene; ds.
XX
OS Alcaligenes faecalis.
XX
PH Key Location/Qualifiers
XX CDS 1..1212
XX FT /*tag= a
XX FT /product= "creatine amidinohydrazide"
XX FT /transl_except= (pos:433..435,aa:Glu)
XX FT /note= "no stop codon given"
XX
PN JP2001346594-A.
XX
PD 18-DEC-2001.
XX
PF 17-DEC-1996; 2001JP-00121708.
XX
PR 17-DEC-1996; 96JP-00337027.
XX
PA (TOYO) TOYOBO KK.
XX
DR WPI: 2002-145187/19.
XX P-PSDB; ABB05660.
XX
PT Mutant creatine amidinohydrazide for use in clinical diagnosis has long-
XX term stability in a neutral buffer solution.
XX
PS Claim 12; Page 12-13; 15pp; Japanese.
XX
CC The present invention describes a stable mutant creatine amidinohydrazide
CC having a long-term stability in a neutral buffer solution. Also described
CC are: (1) a creatine amidinohydrazide gene encoding the above stable
CC creatine amidinohydrazide; (2) a gene encoding a mutant creatine
CC amidinohydrazide having a long-term stability in a neutral buffer solution
CC compared to wild type creatine amidinohydrazide, in which at least one gene
CC among those found in a fully defined 1212 nucleotide sequence (the
CC present sequence) is replaced by another gene; (3) a recombinant plasmid
CC containing a gene encoding the above creatine amidinohydrazide; (4) a cell
CC transformed by the above plasmid; and (5) a method for the preparation of
CC a stable creatine amidinohydrazide in which the above cell is cultured in a
CC medium and creatine amidinohydrazide is collected. The creatine
CC amidinohydrazide is useful as a clinical diagnosing agent
XX
SQ Sequence 1212 BP; 244 A; 400 C; 358 G; 210 T; 0 U; 0 Other;
XX

Query Match 99.9%; Score 1210.4; DB 6; Length 1212;
Best Local Similarity 99.9%; Pred. No. 4.2e-196;
Matches 1211; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ATGCTGACGACATGTTGACGCTGATGAAATGCAACAACGCGGAGAAATATTATTCGCGG 60

```



```

Db      1 ATGACGACGACGATGTTGACGATGATGAAATGGCAACGCGGCAAGAAAATTAATTCGCCG 60
Qy      61 TTTTCGATGCGGAGATGATCCCGCGGCAAAAGACGTTGCGGGCTGATGCGCAAGAC 120
Db      61 TTTTCGATGCGGAGATGATCCCGCGGCAAAAGACGTTGCGGGCTGATGCGCAAGAC 120
Qy      121 AATGTCGATGCGGCGCTGTTTCACTCTTATCACTGATCAACTATTAATTCGCGGTGCTG 180
Db      121 AATGTCGATGCGGCGCTGTTTCACTCTTATCACTGATCAACTATTAATTCGCGGTGCTG 180
Qy      181 TACTGCTATTTGCGAGCGCAAGTACGCGATGTCATGACCAACAACACGCGACAGCATT 240
Db      181 TACTGCTATTTGCGAGCGCAAGTACGCGATGTCATGACCAACAACACGCGACAGCATT 240
Qy      241 TCGGCGGCGCATGACGCGCGGCGCGAGCCCTTGGCGCGCGAGCTTTCGCGGCAACATCACTAC 300
Db      241 TCGGCGGCGCATGACGCGCGGCGCGAGCCCTTGGCGCGCGAGCTTTCGCGGCAACATCACTAC 300
Qy      301 ACCGACTGGCGCGCGCAATTTCTATCGCGCGGTGCGGCACTGACCAACGCGCGCAAG 360
Db      301 ACCGACTGGCGCGCGCAATTTCTATCGCGCGGTGCGGCACTGACCAACGCGCGCAAG 360
Qy      361 CGGATGCGGATCGAGTTGCAACGACGTCATCTGACTTCCCGCGCGACGCTCGAGAAACC 420
Db      361 CGGATGCGGATCGAGTTGCAACGACGTCATCTGACTTCCCGCGCGACGCTCGAGAAACC 420
Qy      421 CTACCGGCGGCTGAGTTGTCGACATGACGCGCGCTCGATGATGATGATGCGCACCATCAAG 480
Db      421 CTACCGGCGGCTGAGTTGTCGACATGACGCGCGCTCGATGATGATGATGCGCACCATCAAG 480
Qy      481 TCGCTCGAAGAGAGAGAGTGTATCCGCGAAGCGCGCGCTGTGTGACGTTCGCGCGCGCG 540
Db      481 TCGCTCGAAGAGAGAGAGTGTATCCGCGAAGCGCGCGCGCTGTGTGACGTTCGCGCGCGCG 540
Qy      541 GCGTGGCGGCGCTGCGCATCAAGCGCGCGGTGCGCGGACGATGAAGTGGGATCGCCACACC 600
Db      541 GCGTGGCGGCGCTGCGCATCAAGCGCGCGGTGCGCGGACGATGAAGTGGGATCGCCACACC 600
Qy      601 AATGCGATGATCGCGGAGATCGCGCAATCGTTCCCTCGTGGAGCTGATGAGCACTGCG 660
Db      601 AATGCGATGATCGCGGAGATCGCGCAATCGTTCCCTCGTGGAGCTGATGAGCACTGCG 660
Qy      661 ACCGCTTCCAGTTCGCGGATCAACACGCGCGCGGACGATCGGTCACCAACCGCATC 720
Db      661 ACCGCTTCCAGTTCGCGGATCAACACGCGCGCGGACGATCGGTCACCAACCGCATC 720
Qy      721 GTGCAATTCGCGGCAATCTTTGCTGCAACACCTTCCGATGATCTTTCGCTACTACCC 780
Db      721 GTGCAATTCGCGGCAATCTTTGCTGCAACACCTTCCGATGATCTTTCGCTACTACCC 780
Qy      781 GCGCTGGAAGGCGACGCTGTTTCTGCGACATGTCGATGACCGCCAGCTTCGATCTGGAG 840
Db      781 GCGCTGGAAGGCGACGCTGTTTCTGCGACATGTCGATGACCGCCAGCTTCGATCTGGAG 840
Qy      841 AAGAAGTGGCGCTGACATCGCGCGGCTGAGCTGATCAAGCGCGCGCGCTGCAAG 900
Db      841 AAGAAGTGGCGCTGACATCGCGCGGCTGAGCTGATCAAGCGCGCGCGCTGCAAG 900
Qy      901 GACATCGCCATCGAGCTCAACGAGATGTACCGGAGTGGAGCTTGTGAATGATCCGCTCC 960
Db      901 GACATCGCCATCGAGCTCAACGAGATGTACCGGAGTGGAGCTTGTGAATGATCCGCTCC 960
Qy      961 TTTGCGCTATGCGCACTCTTTCGCGGTGCTGCTGCACTACGATCGGTGCGGCGGCTG 1020
Db      961 TTTGCGCTATGCGCACTCTTTCGCGGTGCTGCTGCACTACGATCGGTGCGGCGGCTG 1020
Qy      1021 GAGCTGGCGAGGACATCGACACGAGCTGAAGCGCGGAGCTGATGATGATGATGATGATG 1080
Db      1021 GAGCTGGCGAGGACATCGACACGAGCTGAAGCGCGGAGCTGATGATGATGATGATGATG 1080
Qy      1081 AATGTCATGCTGCGGAGGCGATGCGCGGATGCGCGGCGCTATGCGGACGACATCTCTG 1140
Db      1081 AATGTCATGCTGCGGAGGCGATGCGCGGATGCGCGGCGCTATGCGGACGACATCTCTG 1140

```

```

Qy      1141 ATGTCGGGGAGGACGCTGCGCGAAGAACATCAACGCGCTTCCGTTGCTCCGGAACACAC 1200
Db      1141 ATGTCGGGGAGGACGCTGCGCGAAGAACATCAACGCGCTTCCGTTGCTCCGGAACACAC 1200
Qy      1201 ATCATCCGCAAC 1212
Db      1201 ATCATCCGCAAC 1212

RESULT 7
AAT13291
ID      AAT13291 standard; DNA; 1215 BP.
XX
AC      AAT13291;
XX
DT      16-OCT-2003 (revised)
DT      26-SEP-1996 (first entry)
XX
DE      Creatine amidinohydrolyase.
XX
KW      Creatine amidinohydrolyase; CAH; kidney; disease; Alcaligenes; ds.
XX
OS      Alcaligenes sp; KS-85 FERM BP-4487.
XX
FH      Key Location/Qualifiers
FT      misc_difference 319..321
FT      /*tag= a
FT      /tranel_except= 319..321; aa:Aem
XX
XX      DE19536506-A1.
XX
XX      04-APR-1996.
XX
PF      29-SEP-1995; 95DB-01036506.
XX
PR      29-SEP-1994; 94JP-00235737.
XX
XX      (KIRK ) KIRKMAN CORP.
XX
PI      Furukawa K, Ichikawa T, Suzuki M, Koyama Y;
XX
XX      WPI; 1996-180805/19.
XX
DR      P-PSDB; AAR94463.
XX
PT      DNA encoding creatine amidinohydrolyase - useful for quantification of
PT      creatine to, e.g. diagnose kidney disease.
XX
PS      Claim 1; Page 9-11; 18pp; German.
XX
CC      CAH is used for quantification of creatine, e.g. to diagnose kidney
CC      disease by measuring creatine content of serum or urine. CAH DNA can be
CC      inserted into host cells for the prodn. of CAH. CAH can now be produced
CC      efficiently without having to add creatine to the culture medium.
CC      (Updated on 16-OCT-2003 to standardise OS field)
XX
XX      SO
XX      Sequence 1215 BP; 247 A; 398 C; 357 G; 213 T; 0 U; 0 Other;

Query Match      98.0%; Score 1188; DB 2; Length 1215;
Best Local Similarity 98.8%; Pred. No. 2.6e-192;
Matches 1197; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

```

```

Qy      1 ATGACGACGACGATGTTGACGATGATGAAATGGCAACGCGGCAAGAAAATTAATTCGCCG 60
Db      1 ATGACGACGACGATGTTGACGATGATGAAATGGCAACGCGGCAAGAAAATTAATTCGCCG 60
Qy      61 TTTTCGATGCGGAGATGATCCCGCGGCAAAAGACGTTGCGGGCTGATGCGCAAGAC 120
Db      61 TTTTCGATGCGGAGATGATCCCGCGGCAAAAGACGTTGCGGGCTGATGCGCAAGAC 120
Qy      121 AATGTCGATGCGGCGCTGTTTCACTCTTATCACTGATCAACTATTAATTCGCGGTGCTG 180
Db      121 AATGTCGATGCGGCGCTGTTTCACTCTTATCACTGATCAACTATTAATTCGCGGTGCTG 180

```

```

QY 181 TACTGCTATTTGCGAAGCAAGTACGGCATGTGATGACCAACAACAGCCAGAGATT 240
DB 181 TACTGCTATTTGCGAAGCAAGTACGGCATGTGATGACCAACAACAGCCAGCGATT 240
QY 241 TCGGCGGCGCATCGACGGCGGCGAGCCCTGCGCGCGAGCTTGGCGGCAACATCACTAC 300
DB 241 TCGGCGGCGCATCGACGGCGGCGAGCCCTGCGCGCGAGCTTGGCGGCAACATTA 300
QY 301 ACCGACTGCGCGCGCGCAATTTCTATCGCGCGCTGCGCGAGCTGACCAACAGCGCGCAAG 360
DB 301 ACCGACTGCGCGCGCGCAATTTCTATCGCGCGCTGCGCGAGCTGACCAACAGCGCGCAAG 360
QY 361 CGGATGCGGCGCATGATTTGCAACGCTGCAACTGCACTTCCGCGCGGCGAGCTCGAGAAAGC 420
DB 361 CGGATGCGGCGCATGATTTGCAACGCTGCAACTGCACTTCCGCGCGGCGAGCTCGAGAAAGC 420
QY 421 CTACCGGCGCTCGAGTTCTGCAATCAAGCGCAGCCCTGATGATGATGCGGCAACATCAAG 480
DB 421 CTACCGGCGCTCGAGTTCTGCAATCAAGCGCAGCCCTGATGATGATGCGGCAACATCAAG 480
QY 481 TCGCTGAAAGAGAGAAAGCTGATCCGCGAAAGCGCGCGCTGCTGTAAGCTGCGCGCGCG 540
DB 481 TCGCTGAAAGAGAGAAAGCTGATCCGCGAAAGCGCGCGCTGCTGTAAGCTGCGCGCGCG 540
QY 541 GCTGCGCGGCTGCTGATCAAGGCGCGGCTGCGGAGCATGAAATGCGGATCGGCAACAC 600
DB 541 GCTGCGCGGCTGCTGATCAAGGCGCGGCTGCGGAGCATGAAATGCGGATCGGCAACAC 600
QY 601 AATGCGATGATCCGCGAGATCGCAATGCTTCCCTGCTGAGAGCTGATGAGCACTGCG 660
DB 601 AATGCGATGATCCGCGAGATCGCAATGCTTCCCTGCTGAGAGCTGATGAGCACTGCG 660
QY 661 ACCTGCTTCCAGTCGCGGATCAACACGCGCGCGCGCAATCCGCTACCAACCGCATC 720
DB 661 ACCTGCTTCCAGTCGCGGATCAACACGCGCGCGCGCAATCCGCTACCAACCGCATC 720
QY 721 GTGCAATCCGCGGCAATCTTGTGCTCAACCTTCCGATGATCTTCCGCTGCTCAAC 780
DB 721 GTGCAATCCGCGGCAATCTTGTGCTCAACCTTCCGATGATCTTCCGCTGCTCAAC 780
QY 781 GCGCTGAGAGCGAGCTGCTTCTGCGACATGTCGATGATGAGCGCGCTGCACTCTGCGAG 840
DB 781 GCGCTGAGAGCGAGCTGCTTCTGCGACATGTCGATGATGAGCGCGCTGCACTCTGCGAG 840
QY 841 AAGAACGTGCGCGCTGATCGCGCGGCTGAGCTGATCAAGCGCGCGCGCTGCAAG 900
DB 841 AAGAACGTGCGCGCTGATCGCGCGGCTGAGCTGATCAAGCGCGCGCGCTGCAAG 900
QY 901 GACATGCGCATCGAGCTCAACGAGATGTACCGGAGTGGGACCTGCTGAAGTACCGCTCC 960
DB 901 GACATGCGCATCGAGCTCAACGAGATGTACCGGAGTGGGACCTGCTGAAGTACCGCTCC 960
QY 961 TTGCGGATGAGGCACTCTTCCGCGCTGCTGCACTACTACGCTGCGGAGCGCGCGCTG 1020
DB 961 TTGCGGATGAGGCACTCTTCCGCGCTGCTGCACTACTACGCTGCGGAGCGCGCGCTG 1020
QY 1021 GAGCTGCGGAGGCACTCGACACCGAGCTGAGCGCGCGCATGCTTCTCATGAGCGCG 1080
DB 1021 GAGCTGCGGAGGCACTCGACACCGAGCTGAGCGCGCGCATGCTTCTCATGAGCGCG 1080
QY 1081 ATGCTGATGCTGCGGAGGCGCATGCGCGGCTGCGGCGGCTGATGCGGAGCGCATCTCTG 1140
DB 1081 ATGCTGATGCTGCGGAGGCGCATGCGCGGCTGCGGCGGCTGATGCGGAGCGCATCTCTG 1140
QY 1141 ATGCTGAGGAGGAGAGCGGTCGAGGAACATCAACCGGCTTCCGCTGCGGAGCAAC 1200
DB 1141 ATGCTGAGGAGGAGAGCGGTCGAGGAACATCAACCGGCTTCCGCTGCGGAGCAAC 1200
QY 1201 ATCATTCGCGAAC 1212
DB 1201 ATCATTCGCGAAC 1212

```

```

RESULT 8
ACC69514
ID ACC69514 standard; DNA; 1215 BP.
XX
AC ACC69514;
XX
DT 21-JUL-2003 (first entry)
XX
DE Erwinia sp. (DSM 97-934) creatinase encoding DNA SEQ ID NO:1.
XX
KM Erwinia; creatinase; creatine amidinohydrolase; enzyme; EC 3.5.3.3;
XX Erwinia-type creatinase; creatine; creatinine; uraemia; gigantism;
XX chronic nephritis; acute nephritis; tonic muscular dystrophy; gene; ds.
XX Erwinia sp.
OS
XX
FH Key Location/Qualifiers
FT CDS 1..1215
FT /tag= a
FT /EC_number= "3.5.3.3"
FT /Product= "creatinase"
XX
PD EP1298213-A1.
XX
PF 17-SEP-2002; 2002EP-00020793.
XX
PR 20-SEP-2001; 2001EP-00121780.
XX
PA (HOFF) ROCHE DIAGNOSTICS GMBH.
PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
XX
PI Shao Z, Schmuck R, Kratzsch P, Kenklies J, Weisser H;
PI MPI; 2003-383634/37.
XX
DR P-PSDB; ABR43467.
XX
DR
XX
PT New variant of an Erwinia-type creatinase modified relative to a wild-
PT type creatinase having creatinase activity, useful for determining
PT creatinine and/or creatine concentration in a sample.
XX
PS Disclosure; Page 15-17; 51pp; English.
XX
CC The present invention describes a variant of an Erwinia-type creatinase
CC (I) modified relative to a wild-type creatinase having creatinase
CC activity. The variant comprises at least one amino acid substitution at a
CC position of the fully defined 404 amino acid sequence given in ABR43467;
CC these are selected from N130, M203, I278, I304 and F395. Creatinase has
CC the EC number EC 3.5.3.3 and is also known as creatine amidinohydrolase.
CC Also described is a reagent (II) for determining creatine comprising the
CC Erwinia-type creatinase variant. The variant is useful for determining
CC creatinine and/or creatine concentration in a sample. Measuring
CC creatinine and/or creatine are useful for diagnosing uraemia, chronic
CC nephritis, acute nephritis, gigantism, tonic muscular dystrophy and some
CC other related diseases. The mutant enzymes have improved stability, lower
CC conductivity and/or lower Km-values for creatine: they are much better
CC suited to detection methods for creatine. The present sequence encodes
CC Erwinia sp. (DSM 97-934) wild-type creatinase from the present invention
XX
SO Sequence 1215 BP; 242 A; 392 C; 358 G; 223 T; 0 U; 0 Other;
Query Match 81.3%; Score 984.8; DB 8; Length 1215;
Best Local Similarity 88.3%; Pred. No. 6,8e-158;
Matches 1070; Conservative 0; Mismatches 142; Indels 0; Gaps 0;
QY 1 AAGACTGAGAGATGTCAGTCGATGAAATGAGCAAGCGGCGAGAAATTTGCGCG 60
DB 1 AAGACTGAGAGATGTCAGTCGATGAAATGAGCAATGCGAGAAATTTTCCCCC 60
QY 61 TTTTGGATGCGGAGATGACCGCGCGCAAAAGAGCTTGGCGCTGATGCGCAAGAAC 120
DB 61 TTTTGGATGCGGAGATGACCGCGCGCAAAAGAGCTTGGCGCTGATGCGCAAGAAC 120

```

```

Db      61  TTTCCGATGCCAGATGACGCCGCCGACAGATGACGTGGCGCTGATGCGCGAAG 120
Qy      121  AATGTCATGCGCGCGCTGTTCACTCTTATCACTGATCACTACTATTCCGCTGCTG 180
Db      121  GACGTGACGCTGCGCTGTTCACTCTTATCACTGATCACTACTATTCTGATTCCTG 180
Qy      181  TACTGCTATTTCGACGCAAGTACCGGATGATGACCAACAACCGCACGACGATT 240
Db      181  TACTGCTATTTCGCGCAATACCGGATGATGACCAACAACCGCACGACGATT 240
Qy      241  TCGCGCGCATCGACCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
Db      241  TCGCGCGCATCGATCGCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
Qy      301  ACCGATCGCGCGCGCGCAATTTCTATCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
Db      301  ACCGATCGCGCGCGCGCAATTTCTATCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
Qy      361  CGCATCGCGCATCGATTCGACACGTCATCTGACCTTCGCGCGCGCGCGCGCGCG 420
Db      361  CGCATCGCGATCGAGTTTCGACACGTCATCTGACCTTCGCGCGCGCGCGCGCG 420
Qy      421  CTACCGCGCGCTCGAGTTCTGTCATCATGACCGCGCTCGATGATGATGCGCAC 480
Db      421  CTGCGCGCGCTCGAGTTCTGTCATCATGATCGATCGATCGATGATGATGCGCA 480
Qy      481  TCGCTCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Db      481  TCGCTCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Qy      541  GCGTCGCGCGCTGCGCATCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
Db      541  GCGTCGCTGCGCGCGCATCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
Qy      601  AATGCGATGATCGCGCATGATCGCGCATGATGATGATGATGATGATGATGATG 660
Db      601  AATGCGATGATCGCGCATGATCGCGCATGATGATGATGATGATGATGATGATG 660
Qy      661  ACTGCTTCATGCGCGCGCATCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
Db      661  ACTGCTTCATGCGCGCGCATCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
Qy      721  GTGCAATCGCGCGCATCTCTTCTGCTCAACACTTCCCGATGATCTTCCGCTAC 780
Db      721  GTGCAATCGCGCGCATCTCTTCTGCTCAACACTTCCCGATGATCTTCCGCTAC 780
Qy      781  GCGCTGAGACGCACTGCTGCTGCGCACTGTCGATGATGATGATGATGATGATG 840
Db      781  GCGCTGAGACGCACTGCTGCTGCGCACTGTCGATGATGATGATGATGATGATG 840
Qy      841  AAGAACGTGCGCGTGCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
Db      841  AAGAACGTGCGCGTGCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
Qy      901  GACATCGCATGAGCTCAACGAGATGATCCGCGAGTGGGACTGTCGAAGTACG 960
Db      901  GATATGCGCATGAGCTCAACGAGATGATCCGCGAGTGGGACTGTCGAAGTACG 960
Qy      961  TTGCGCTATGCGCACTCTTCTGCGCTGCTGTCGATGATGATGATGATGATG 1020
Db      961  TTGCGCTATGCGCACTCTTCTGCGCTGCTGTCGATGATGATGATGATGATG 1020
Qy      1021  GAGCTGCGCGAGGACATCGACACCGAGCTGAGAGCGCGCGCGCGCGCGCGCG 1080
Db      1021  GAGCTGCGCGAGGACATCGACACCGAGCTGAGAGCGCGCGCGCGCGCGCGCG 1080
Qy      1081  AATGCGATGCTGCGCGAGGAGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1140
Db      1081  AATGCGATGCTGCGCGAGGAGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1140
Qy      1141  ATGCTGCGCGAGGACGCTGCGCGAGGACATCAACCGCGCTTCCGCTGCGGAA 1200
Db      1141  ATGCTGCGCGAGGACGCTGCGCGAGGACATCAACCGCGCTTCCGCTGCGGAA 1200

```

```

Qy      1201  ATCATCCGCAAC 1212
Db      1201  ATCATCCGCAAC 1212

RESULT 9
ACC69519
ID ACC69519 standard; DNA; 1215 BP.
XX
AC ACC69519;
XX
DT 21-JUL-2003 (first entry)
XX
DE Mutant Erwinia creatinase CTgc2 encoding DNA SEQ ID NO:17.
XX
KW Erwinia; creatinase; creatine amidinohydrolase; enzyme; EC 3.5.3.3;
KW Erwinia-type creatinase; creatine; creatinine; uraemia; gigantism;
KW chronic nephritis; acute nephritis; tonic muscular dystrophy; mutant;
KW gene; ds.
XX
OS Erwinia sp.
XX Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 1..1215
FT /tag= a
FT /EC_number= "3.5.3.3"
FT /product= "creatinase mutant CTgc2"
XX
PN EP1298213-A1.
XX
PD 02-APR-2003.
XX
PF 17-SEP-2002; 2002EP-00020793.
XX
PR 20-SEP-2001; 2001EP-00121780.
XX
PA (HOPE ) ROCHE DIAGNOSTICS GMBH.
PA (HOPE ) HOFFMANN LA ROCHE & CO AG F.
XX
PI Shao Z, Schmuck R, Kratzsch P, Kenkies J, Weisner H;
XX
DR MPI; 2003-383834/37.
XX
DR P-PDB; ABR43478.
XX
PT New variant of an Erwinia-type creatinase modified relative to a wild-
PT type creatinase having creatinase activity, useful for determining
PT creatinine and/or creatine concentration in a sample.
XX
PS Example 4; Page 33-35; 51pp; English.
XX
CC The present invention describes a variant of an Erwinia-type creatinase
CC (I) modified relative to a wild-type creatinase having creatinase
CC activity. The variant comprises at least one amino acid substitution at a
CC position of the fully defined 404 amino acid sequence given in ABR43467;
CC these are selected from N130, M203, I278, I1304 and P195. Creatinase has
CC the EC number EC 3.5.3.3 and is also known as creatine amidinohydrolase.
CC Also described is a reagent (II) for determining creatine comprising the
CC Erwinia-type creatinase variant. The variant is useful for determining
CC creatinine and/or creatine concentration in a sample. Measuring
CC creatinine and creatine are useful for diagnosing uraemia, chronic
CC nephritis, acute nephritis, gigantism, tonic muscular dystrophy and some
CC other related diseases. The mutant enzymes have improved stability, lower
CC conductivity and/or lower Km-values for creatine; they are much better
CC suited to detection methods for creatine. The present sequence encodes a
CC mutant Erwinia creatinase from the present invention
XX
SO Sequence 1215 BP; 241 A; 392 C; 361 G; 221 T; 0 U; 0 Other;

Query Match 81.3%; Score 984.8; DB 8; Length 1215;
Best Local Similarity 88.3%; Pred. No. 6,8e-158;
Matches 1070; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

```

QY 1 ATGACTGACGACGATGTTGACGCTGATGAATGACACAGCGGAGAAATATTGCGCG 60
 Db 1 ATGACTGACGACGATGTTGACGCTGATGAATGACACAGCGGAGAAATATTGCGCG 60
 QY 61 TTTTGGAGATGCGGAGATGACCCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
 Db 61 TTTTGGAGATGCGGAGATGACCCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
 QY 121 AATGTCGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
 Db 121 AATGTCGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
 QY 181 TACTGCTATTTGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 240
 Db 181 TACTGCTATTTGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 240
 QY 241 TGGGCGGCGGATGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
 Db 241 TGGGCGGCGGATGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
 QY 301 ACCGACTGCG 360
 Db 301 ACCGACTGCG 360
 QY 361 GCGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 420
 Db 361 GCGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 420
 QY 421 CTACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
 Db 421 CTACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
 QY 481 TGGCTGGAAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 540
 Db 481 TGGCTGGAAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 540
 QY 541 GCGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 600
 Db 541 GCGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 600
 QY 601 AATGCGATGATCGCGGAGATCGCGAAATGCTTCCCTTGGAGTGAAGACGACCTGAG 660
 Db 601 AATGCGATGATCGCGGAGATCGCGAAATGCTTCCCTTGGAGTGAAGACGACCTGAG 660
 QY 661 ACCGCTTTCAGTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 720
 Db 661 ACCGCTTTCAGTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 720
 QY 721 GTGCAATCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 780
 Db 721 GTGCAATCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 780
 QY 781 GGGCTGGAAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 840
 Db 781 GGGCTGGAAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 840
 QY 841 AAGAACTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 900
 Db 841 AAGAACTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 900
 QY 901 GACATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960
 Db 901 GACATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960
 QY 961 TTTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1020
 Db 961 TTTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1020
 QY 1021 GAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080
 Db 1021 GAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080

QY 1081 ATGCGATGCTGCGGAGAGGAGATGCGGAGTCCGCGGCGGCGGCTATTCGCGAGACGACATCC 1140
 Db 1081 ATGCGATGCTGCGGAGAGGAGATGCGGAGTCCGCGGCGGCGGCTATTCGCGAGACGACATCC 1140
 QY 1141 ATTCGTCGCGGAGAGAGCGTCCGAGAGACATCACCGGCTTCCGCTTGGCTCCGAGACAAAC 1200
 Db 1141 ATTCGTCGCGGAGAGAGCGGCGGAGAGACATTCACCGATTCCTTCCGCTTGGCTCCGAGACAAAC 1200
 QY 1201 ATCATCCGCAAC 1212
 Db 1201 ATCATCCGCAAC 1212
 RESULT 10
 ACC69515
 ID ACC69515 standard; DNA; 1215 BP.
 XX
 AC C69515;
 XX
 DT 21-JUL-2003 (first entry)
 XX
 DE Mutant Erwinia creatinase CTIm24 encoding DNA SEQ ID NO:9.
 XX
 KW Erwinia: creatinase; creatine amidinohydrolase; enzyme; EC 3.5.3.3;
 KW Erwinia-type creatinase; creatine; creatinine; uremia; gigantism;
 KW chronic nephritis; acute nephritis; tonic muscular dystrophy; mutant;
 KW gene; ds.
 XX
 OS Erwinia sp.
 OS Synthetic.
 FH
 FT Key Location/Qualifiers
 FT CDS 1..1215
 FT /*tag= a
 FT /EC number= "3.5.3.3"
 FT /product= "creatinase mutant CTIm24"
 XX
 PN EP1298213-A1.
 XX
 PD 02-APR-2003.
 XX
 PF 17-SEP-2002; 2002EP-00020793.
 XX
 PR 20-SEP-2001; 2001EP-00121780.
 XX
 PA (HOFF) ROCHE DIAGNOSTICS GMBH.
 PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
 XX
 PI Shao Z, Schmuck R, Kratzsch P, Kenkles J, Weisner H;
 DR WPI; 2003-383834/37.
 DR P-PSDB; ABR43474.
 XX
 PT New variant of an Erwinia-type creatinase modified relative to a wild-
 PT type creatinase having creatinase activity, useful for determining
 PT creatinine and/or creatine concentration in a sample.
 XX
 PS Example 4; Page 20-21; 51pp; English.
 XX
 CC The present invention describes a variant of an Erwinia-type creatinase
 CC (I) modified relative to a wild-type creatinase having creatinase
 CC activity. The variant comprises at least one amino acid substitution at a
 CC position of the fully defined 404 amino acid sequence given in ABR43467;
 CC these are selected from N130, M203, I278, I1304 and F395. Creatinase has
 CC the EC number EC 3.5.3.3 and is also known as creatine amidinohydrolase.
 CC Also described is a reagent (II) for determining creatine comprising the
 CC Erwinia-type creatinase variant. The variant is useful for determining
 CC creatinine and/or creatine concentration in a sample. Measuring
 CC creatinine and creatine are useful for diagnosing uraemia, chronic
 CC nephritis, acute nephritis, gigantism, tonic muscular dystrophy and some
 CC other related diseases. The mutant enzymes have improved stability, lower
 CC conductivity and/or lower Km-values for creatine; they are much better
 CC suited to detection methods for creatine. The present sequence encodes a

CC mutant Erwinia creatinase from the present invention
 XX Sequence 1215 BP; 241 A; 393 C; 359 G; 222 T; 0 U; 0 Other:
 Query Match 81.0%; Score 981.6; DB 8; Length 1215;
 Best Local Similarity 88.1%; Pred. No. 2,4e-157;
 Matches 1068; Conservative 0; Mismatches 144; Indels 0; Gaps 0;

QY 1 ATGACGAGAGCAATGTTGACGATGAATGGCAACAGGAGAGAAATTTTGGCG 60
 DB 1 ATGACGAGAGCAATGTTGACGATGAATGGCAACAGGAGAGAAATTTTGGCG 60
 QY 61 TTTTCGATGCCAGATGACCCGCGCAAAAGAGTTCCGCGCTGATGCGCAAGAC 120
 DB 61 TTTTCGATGCCAGATGACCCGCGCAAAAGAGTTCCGCGCTGATGCGCAAGAC 120
 QY 121 AATGTCGATGCGGCGCTGTTACCTTATCATGTCATCACTAATTCCGCTGAG 180
 DB 121 GACGTGACGCTGCGCTGTTACCTTATCATGTCATCACTAATTCCGCTGAG 180
 QY 181 TACTGCTATTTGGAGCAAGTACGGCATGTCATGACCAAGCAACGCGACAGATT 240
 DB 181 TACTGCTATTTGGAGCAAGTACGGCATGTCATGACCAAGCAACGCGACAGATT 240
 QY 241 TCGGCGCGGATCGACGCGCGCGACGCTGCGCGCGCGACGCTGCGCGCAACATCACTAC 300
 DB 241 TCGGCGCGGATCGACGCGCGCGCGACGCTGCGCGCGCGACGCTGCGCGCAACATCACTAC 300
 QY 301 ACCGATCGGCGCGCGCGCAATTTCTATCGCGCGCGCGCGCGCACTGACAGGCGCCAG 360
 DB 301 ACCGATCGGCGCGCGCGCAATTTCTATCGCGCGCGCGCGCGCACTGACAGGCGCCAG 360
 QY 361 CGCATGCGATGAGATTTCACACCGTCATCTGCACTTCGCGCGCGCGCGCAAGGCG 420
 DB 361 CGCATGCGATGAGATTTCACACCGTCATCTGCACTTCGCGCGCGCGCGCAAGGCG 420
 QY 421 CTACGCGCGGCTGATGCTGTCGATCAGCGACGCTGATGATGATGCGCAACATCAAG 480
 DB 421 CTACGCGCGGCTGATGCTGTCGATCAGCGACGCTGATGATGATGCGCAACATCAAG 480
 QY 481 TCGCTGGAAGAGAGAGAGCTGATCGCGAGGCGCGCGCGCGCGCGCGCGCGCG 540
 DB 481 TCGCTGGAAGAGAGAGAGCTGATCGCGAGGCGCGCGCGCGCGCGCGCGCGCG 540
 QY 541 GCGTGGCG 600
 DB 541 GCGTGGCG 600
 QY 601 AATGCGATGATCGCGAGATCGCGCAATGTTCCCTGCTGAGGCTGATGAGCACTGCG 660
 DB 601 AATGCGATGATCGCGAGATCGCGCAATGTTCCCTGCTGAGGCTGATGAGCACTGCG 660
 QY 661 ACCTGGTTCCAGTCGCGCGCATCAACACGCGCGCGCGCGCGCGCGCGCGCGCATC 720
 DB 661 ACCTGGTTCCAGTCGCGCGCATCAACACGCGCGCGCGCGCGCGCGCGCGCATC 720
 QY 721 GTGCAATCCGCGGACATCTCTTGGCTCAACCTTCCGATGATCTTGGCTGATCAAC 780
 DB 721 GTGCAATCCGCGGACATCTCTTGGCTCAACCTTCCGATGATCTTGGCTGATCAAC 780
 QY 781 GGGCTGAGAGCGAGCGCTGTTGCGACGATGATGAGGCGCGCTGACATCTGAGAG 840
 DB 781 GGGCTGAGAGCGAGCGCTGTTGCGACGATGATGAGGCGCGCTGACATCTGAGAG 840
 QY 841 AAGAAAGTGGCGGTGATGCGCGCGCGCTGAGCTGATCAAGCGCGCGCGCGCTGCAAG 900
 DB 841 AAGAAAGTGGCGGTGATGCGCGCGCGCTGAGCTGATCAAGCGCGCGCGCGCTGCAAG 900
 QY 901 GACATGCGCATGAGCTCAACGAGATGATCCGCGATGAGGAGCTGCTGAGTACCGCTCC 960
 DB 901 GATATCGCATGAGCTCAACGAGATGATCCGCGATGAGGAGCTGCTGAGTACCGCTCC 960
 QY 961 TTGCGGTATGAGCACTCTTGGCGCGGTGCTGTCACACTACGAGGTGCGAGGCGCGCGT 1020

DB 961 TTGCGGTATGAGCACTCTTGGCGCGGTGCTGTCACACTACGAGGTGCGAGGCGCGCGT 1020
 QY 1021 GAGCTGGCGGAGAGACATCGACACCGAGCTGAGAGCCCGCGAGATGCTGCTCCATGAGCG 1080
 DB 1021 GAGCTGGCGGAGAGACATCGATACCGGCTGAGAGCCCGCGAGATGCTGCTCCATGAGCG 1080
 QY 1081 ATGATGATCTCTCCGAGAGGAGCATGCGCGGTGCGCGCGCTATGCGAGACGACATCTG 1140
 DB 1081 ATGATGATCTCTCCGAGAGGAGCATGCGCGGTGCGCGCGCTATGCGAGACGACATCTG 1140
 QY 1141 ATGCTGGGAGAGACGCTGCGAGAACATCACCGCTTCCGTTCCGTTCCGAAACACAC 1200
 DB 1141 ATGCTGGGAGAGACGCTGCGAGAACATCACCGCTTCCGTTCCGTTCCGAAACACAC 1200
 QY 1201 ATCATCCGCAAC 1212
 DB 1201 ATCATCCGCAAC 1212

RESULT 11
 ACC69517
 ID ACC69517 standard; DNA, 1212 BP.
 AC ACC69517;
 XX 21-JUL-2003 (first entry)
 XX
 XX
 DB Mutant Erwinia creatinase CT2m10 encoding DNA SEQ ID NO:13.
 XX
 XX
 XX Erwinia; creatinase; creatine amidinohydrolase; enzyme; EC 3.5.3.3;
 KW Erwinia-type creatinase; creatine; creatinine; uraemia; gigantism;
 KW chronic nephritis; acute nephritis; tonic muscular dystrophy; mutant;
 KW gene; de.
 XX
 OS Erwinia sp.
 OS Synthetic.
 XX
 XX
 FT Key Location/Qualifiers
 FT CDS 1..1212
 FT /tag= a
 FT /partial
 FT /EC number= "3.5.3.3"
 FT /product= "creatinase mutant CT2m10"
 FT /note= "no stop codon given"

PN EP1298213-A1.
 XX
 PD 02-APR-2003.
 XX
 PF 17-SEP-2002; 2002EP-00020793.
 XX
 PR 20-SEP-2001; 2001EP-00121780.
 XX
 PA (HOPE) ROCHE DIAGNOSTICS GMBH.
 PA (HOPE) HOFFMANN LA ROCHE & CO AG F.
 XX
 PI Shao Z, Schmuck R, Kratzsch P, Kenkies J, Weisser H;
 XX
 DR MPI; 2003-38384/37.
 DR P-PSDB; ABR43476.
 XX
 PT New variant of an Erwinia-type creatinase modified relative to a wild-
 PT type creatinase having creatinase activity, useful for determining
 PT creatinine and/or creatine concentration in a sample.
 XX
 PS Example 4; Page 26-28; 51pp; English.
 XX
 CC The present invention describes a variant of an Erwinia-type creatinase
 CC (1) modified relative to a wild-type creatinase having creatinase
 CC activity. The variant comprises at least one amino acid substitution at a
 CC position of the fully defined 404 amino acid sequence given in ABR43467;
 CC these are selected from N130, M203, I278, I1304 and F395. Creatinase has

CC the EC number EC 3.5.3.3 and is also known as creatine amidinohydrolase.
 CC Also described is a reagent (II) for determining creatine comprising the
 CC Erwinia-type creatinase variant. The variant is useful for determining
 CC creatinine and/or creatine concentration in a sample. Measuring
 CC creatinine and creatine are useful for diagnosing uraemia, chronic
 CC nephritis, acute nephritis, gigantism, tonic muscular dystrophy and some
 CC other related diseases. The mutant enzymes have improved stability, lower
 CC conductivity and/or lower Km-values for creatine; they are much better
 CC suited to detection methods for creatine. The present sequence encodes a
 CC mutant Erwinia creatinase from the present invention
 CC

Sequence 1212 BP; 240 A; 394 C; 358 G; 220 T; 0 U; 0 Other;

Query Match 80.3%; Score 980; DB 8; Length 1212;
 Best Local Similarity 88.0%; Pred. No. 4,4e-157;
 Matches 1067; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

QY 1 ATGACTGACGACATGTTGACGATGATGAATGGACAAACGCGAGAAAGATTATTCGCG 60
 DB 1 ATGACTGACGACATGTTGACGATGATGAATGGACAAATGTAAGAAATATTCGCG 60
 QY 61 TTTTCGATGCGGACGATGACCCGCGCAAAACGACGTTGCGCGTGGATGGCCAAAC 120
 DB 61 TTTTCGATGCGGACGATGACCCGCGCAAGCGCGGACGATGACGCGTGGATGGCCAAAC 120
 QY 121 AATGTGATGCGGCGCTGTTCACTTATCACTGATCACTAATCTATTCGCGTGGCTG 180
 DB 121 GACGTGACGCTGCGCTGTTCACTTATCACTGATCACTAATCTATTCGCGTGGCTG 180
 QY 181 TACTGCTATTGGAGACGACAGTACGGGATGATGACGACCAACACGCGACGACGATT 240
 DB 181 TACTGCTATTGGAGACGACAGTACGGGATGATGACGACCAACACGCGACGACGATT 240
 QY 241 TCGGCGCGGACATGACGCGCGGACGCGCTGCGCGGACGCTTGGCGACCAATCACTAC 300
 DB 241 TCGGCGCGGACATGACGCGCGGACGCGCTGCGCGGACGCTTGGCGACCAATCACTAC 300
 QY 301 ACCGACTGCGCGCGCGGACCAATTTCTATCGCGCGCTGCGCGGACGCTTGGCGACCA 360
 DB 301 ACCGACTGCGCGCGCGGACCAATTTCTATCGCGCGCTGCGCGGACGCTTGGCGACCA 360
 QY 361 CGCATGCGGATCGAGTTGACGACGATCACTTCCGCGCGGACGCTTGGCGACCAAG 420
 DB 361 CGCATGCGGATCGAGTTGACGACGATCACTTCCGCGCGGACGCTTGGCGACCAAG 420
 QY 421 CTACCGGCGCTGACGTTGATGACATGACGACGCGCTGATGATGACGACATCAAG 480
 DB 421 CTACCGGCGCTGACGTTGATGACATGACGACGCGCTGATGATGACGACATCAAG 480
 QY 481 TCGCTGGAAGACGAAAGCTGATCCGCGAAGGCGCGCGTGTGTGACGTCGCGCGGCG 540
 DB 481 TCGCTGGAAGACGAAAGCTGATCCGCGAAGGCGCGCGTGTGTGACGTCGCGCGGCG 540
 QY 541 GCGTGGCGGCTGCGATCAAGCGCGGCGGCGCGGATGAGTGGGATCGCACCAAC 600
 DB 541 GCGTGGCGGCTGCGATCAAGCGCGGCGGCGGATGAGTGGGATCGCACCAAC 600
 QY 601 AATGCAATGATCGCGGATCGCAAAATCGTTCCCTTGTGATGATGACCACTGG 660
 DB 601 AATGCAATGATCGCGGATCGCAAAATCGTTCCCTTGTGATGATGACCACTGG 660
 QY 661 ACCGTGTTCAAGTGGGATCAACACGAGCGCGGACCAATCGGATGACCAACCGGATC 720
 DB 661 ACCGTGTTCAAGTGGGATCAACACGAGCGCGGACCAATCGGATGACCAACCGGATC 720
 QY 721 GTGCAATCGGCGGATCGTTGCTGCTCAACGTTCCGATGATCTTCGCTACTACACG 780
 DB 721 GTGCAATCGGCGGATCGTTGCTGCTCAACGTTCCGATGATCTTCGCTACTACACG 780
 QY 781 GCGCTGAGGACGACGCTGTTCTGCGACCATGCGATGACGACGCTGACATCTGGAG 840
 DB 781 GCGCTGAGGACGACGCTGTTCTGCGACCATGCGATGACGACGCTGACATCTGGAG 840

QY 841 AAGACGTGCGCGGTGATCGCGCGGCTCGAGCTGATCAAGCGCGGCGCTGCAAG 900
 DB 841 AAGACGTGCGCGGTGATCGCGCGGCTCGAGCTGATCAAGCGCGGCTGCGCTCAAG 900
 QY 901 GACATGCGCATGAGCTCAACGATGTAACGCGAGTGGGACCTGCTGAATGCCCTCC 960
 DB 901 GATATGCGCATGAGCTCAACGATGTAACGCGAGTGGGACCTGCTGAATGCCCTCC 960
 QY 961 TTGCGTATGCGGACCTCTTCCGCGCTGCTGCTGCTCACTACTACGCTGCGAGCGGCG 1020
 DB 961 TTGCGTATGCGGACCTCTTCCGCGCTGCTGCTGCTCACTACTACGCTGCGAGCGGCG 1020
 QY 1021 GAGCTGCGGAGGACATGACACCGAGCTGAAGCCGCGGATGCTTCTCATGAGCG 1080
 DB 1021 GAGCTGCGGAGGACATGACATTCAGTCTGACGCGCGGATGCTTCTCATGAGCG 1080
 QY 1081 ATGCTGATGCTGCGGAGGACATGCGCGGCTGCGGCTGCTATGCGACGACATCTTG 1140
 DB 1081 ATGCTGATGCTGCGGAGGACATGCGCGGCTGCGGCTGCTATGCGACGACATCTTG 1140
 QY 1141 ATGCTGCGGAGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
 DB 1141 ATGCTGCGGAGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
 QY 1201 ATGATCGGCAAC 1212
 DB 1201 ATGATCGGCAAC 1212

RESULT 12
 ID ACC69516
 ACC69516 standard; DNA; 1212 BP.
 XX AC ACC69516;
 XX DT 21-JUL-2003 (first entry)
 XX DE Mutant Erwinia creatinase CT2m9 encoding DNA SEQ ID NO.11.
 KW Erwinia; creatinase; creatine amidinohydrolase; enzyme; EC 3.5.3.3;
 KW Erwinia-type creatinase; creatine; creatinine; uraemia; gigantism;
 KW chronic nephritis; acute nephritis; tonic muscular dystrophy; mutant;
 KM gene; de.
 OS Erwinia sp.
 OS Synthetic.
 OS
 FH Key Location/Qualifiers
 FT CDS 1..1212
 FT /tag= a
 FT /partial
 FT /EC_number= "3.5.3.3"
 FT /product= "creatinase mutant CT2m9"
 FT /note= "no stop codon given"
 XX
 PN EP1298213-A1.
 XX
 PD 02-APR-2003.
 XX
 PF 17-SEP-2002; 2002BP-00020793.
 XX
 PR 20-SEP-2001; 2001BP-00121780.
 XX
 PA (HOFF) ROCHE DIAGNOSTICS GMBH.
 PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
 XX
 PI Shao Z, Schmuck R, Kratzsch P, Kenklee J, Weisser H;
 DR WPI; 2003-383834/37.
 DR P-PDB; ABR3475.
 XX
 PT New variant of an Erwinia-type creatinase modified relative to a wild-type creatinase having creatinase activity, useful for determining

PT creatinine and/or creatine concentration in a sample.

XX Example 4; Page 23-25; 51pp; English.

CC The present invention describes a variant of an Erwinia-type creatinase
 CC (1) modified relative to a wild-type creatinase having creatinase
 CC activity. The variant comprises at least one amino acid substitution at a
 CC position of the fully defined 404 amino acid sequence given in ABR43467;
 CC these are selected from N130, M203, I278, I1304 and P395. Creatinase has
 CC the EC number EC 3.5.3.3 and is also known as creatine amidinohydrolase.
 CC Also described is a reagent (II) for determining creatine comprising the
 CC Erwinia-type creatinase variant. The variant is useful for determining
 CC creatinine and/or creatine concentration in a sample. Measuring
 CC creatinine and creatine are useful for diagnosing uraemia, chronic
 CC nephritis, acute nephritis, gigantism, tonic muscular dystrophy and some
 CC other related diseases. The mutant enzymes have improved stability, lower
 CC conductivity and/or lower Km-values for creatine; they are much better
 CC suited to detection methods for creatine. The present sequence encodes a
 CC mutant Erwinia creatinase from the present invention

XX Sequence 1212 BP; 239 A; 393 C; 359 G; 221 T; 0 U; 0 Other;

Query Match 80.9%; Score 980; DB 8; Length 1212;

Best Local Similarity 88.0%; Pred. No. 4,4e-157;

Matches 1067; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

QY 1 ATGACGTGAGCAATGTTGACGTGATGAAATGGACACAGCGGAGAAAGATTATTTGGCGG 60
 DB 1 ATGACGTGAGCAATGTTGACGTGATGAAATGGACACAGCGGAGAAATTTCCCGC 60
 QY 61 TTTTCGGATGCCGAGATGACCCCGCCCAAAACGACCTTGGCGGATGAGCCCAAGAC 120
 DB 61 TTTTCGGATGCCGAGATGACCCCGCCCAAGATGACCTTGGCGGATGAGCCCAAGAC 120
 QY 121 AATGTCGATGCGGCGCTGTTACCTCTTATCACTGATCACTACTATTTCCGCTGGCTG 180
 DB 121 GACGTGACGCTGCGCTGTTACCTCTTATCACTGATCACTACTATTTCCGCTGGCTG 180
 QY 181 TACTGCTATTTGGAGAGCAAGTACGGGATGATGATGACACAAACGCGACAGCAT 240
 DB 181 TACTGCTATTTGGAGAGCAATACGGGATGATGATGACACAAACGCGACAGCAT 240
 QY 241 TCGGCGCGGATCGACCGGCGGACGCTTGGCGGCGGACGTTGGCGGACATCACTAC 300
 DB 241 TCGGCGCGGATCGATGCGGCGGACGCTTGGCGGCGGACGTTGGCGGACATCACTAC 300
 QY 301 ACCGACTGCGGCGGCGGACATTTCTATCGGCGGCTGCGGCGGATGACACAGCGGCG 360
 DB 301 ACCGACTGCGGCGGCGGACATTTCTATCGGCGGCTGCGGCGGATGACACAGCGGCG 360
 QY 361 CGGATGCGGATCGATGCGGCGGACGCTTGAATCTGATCTTCCGCGGCGGATGAGAAC 420
 DB 361 CGGATGCGGATCGATGCGGCGGACGCTTGAATCTGATCTTCCGCGGCGGATGAGAAC 420
 QY 421 CTACGCGGCGGATGCGGCGGACGCTTGAATCTGATCTTCCGCGGCGGATGAGAAC 480
 DB 421 CTACGCGGCGGATGCGGCGGACGCTTGAATCTGATCTTCCGCGGCGGATGAGAAC 480
 QY 481 TCGCTCGAAGAGCAAGTATCGGCGGAGCGCGCGCTGTGTGACGTGCGGCGGCG 540
 DB 481 TCGCTCGAAGAGCAAGTATCGGCGGAGCGCGCGCTGTGTGACGTGCGGCGGCGGCG 540
 QY 541 GCGTGGCGGCTGCGATCAAGCGGCGGCTGCGGAGCATGAGTGGCGATGCGCACACC 600
 DB 541 GCGTGGCGGCTGCGATCAAGCGGCGGCTGCGGAGCATGAGTGGCGATGCGCACACC 600
 QY 601 AATGCGATGATCGGCGGAGATCGCAATGTTCCCTTGGAGGATGAGCACTGG 660
 DB 601 AATGCGATGATCGGCGGAGATCGCAATGTTCCCTTGGAGGATGAGCACTGG 660
 QY 661 ACTGCTTCAGTGGGCGATCAACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720
 DB 661 ACTGCTTCAGTGGGCGATCAACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720

QY 721 GTGCAATCGGCGGAGCATCTTTGCTCAACACCTTCCGATGATCTTGGCTACTACACC 780
 DB 721 GTGCAATCGGCGGAGCATCTTTGCTCAACACCTTCCGATGATCTTGGCTACTACACC 780
 QY 781 GCGCTGAGAGGACGCTGTTTTCGACACATGTCGATGAGCGGACCTGCAATCTGGAG 840
 DB 781 GCGCTGAGAGGACGCTGTTTTCGACACATGTCGATGAGCGGACCTTGAACCTGGAG 840
 QY 841 AAGAACGTGCGGCTGATGCGGCGGCTGACCTGATCAAGCGGCGGCGGCTGCAAG 900
 DB 841 AAGAACGTGCGGCTGATGCGGCGGCTGACCTGATCAAGCGGCGGCGGCTGCAAG 900
 QY 901 GACATGCGCATCGAGCTCAACAGATGTAACCGGAGTGGGACCTGCTGAAGTACCGCTCC 960
 DB 901 GATATGCGCATCGAGCTCAACAGATGTAACCGGAGTGGGATGCTGTAAGTACCGCTCC 960
 QY 961 TTCGCTATGCGCATCTTCTTGGCGGCTGTCGCTGCTACTACTAAGTGGCGGCGGCTG 1020
 DB 961 TTCGCTATGCGCATCTTCTTGGCGGCTGTCGCTGCTACTACTAAGTGGCGGCGGCTG 1020
 QY 1021 GAGCTGCGGAGGACATCGACACCGGCTGAAAGCGGCGATGAGTGTCTCATGAGCGG 1080
 DB 1021 GAGCTGCGGAGGACATCGACACCGGCTGAAAGCGGCGATGAGTGTCTCATGAGCGG 1080
 QY 1081 ATGCTGATGCTGCGGAGGACATGCGGCTGCGGCGGCTATGCGGACGACATCTCTG 1140
 DB 1081 ATGCTGATGCTGCGGAGGACATGCGGCTGCGGCGGCTATGCGGACGACATCTCTG 1140
 QY 1141 ATGCTGCGGAGGAGACGCTGCGGAGGACATGCGGCTTCCGCTTGGCTGCGGAGCAAC 1200
 DB 1141 ATGCTGCGGAGGAGACGCTGCGGAGGACATGCGGCTTCCGCTTGGCTGCGGAGCAAC 1200
 QY 1201 ATCATCGGCAAC 1212
 DB 1201 ATCATCGGCAAC 1212

RESULT 13
 ACCG9518
 ID ACCG9518 standard; DNA; 1215 BP.
 XX AC ACCG9518;
 XX DT 21-JUL-2003 (first entry)
 XX DE Mutant Erwinia creatinase CT2m28 encoding DNA SEQ ID NO:15.
 XX KW Erwinia; creatinase; creatine amidinohydrolase; enzyme; EC 3.5.3.3;
 XX KW Erwinia-type creatinase; creatine; creatinine; uraemia; gigantism;
 XX KW chronic nephritis; acute nephritis; tonic muscular dystrophy; mutant;
 XX KW gene; ds.
 XX OS Erwinia sp.
 XX OS Synthetic.
 XX FH Key
 XX FT CDS
 XX FT Location/Qualifiers
 XX FT 1..1215
 XX FT /tag= a
 XX FT /EC number= "3.5.3.3"
 XX FT /product= "creatinase mutant CT2m28"
 XX PN BP1298213-A1.
 XX PD 02-APR-2003.
 XX PD 17-SEP-2002; 2002BP-00020793.
 XX PR 20-SEP-2001; 2001BP-00121780.
 XX PA (HOFF) ROCHE DIAGNOSTICS GMBH.
 XX PA (HOFF) HOFFMANN LA ROCHE & CO AG F.

PD 02-APR-2003.
 XX
 PF 17-SEP-2002; 2002EP-00020793.
 XX
 PR 20-SEP-2001; 2001EP-00121780.
 XX
 PA (HOPE) ROCHE DIAGNOSTICS GMBH.
 PA (HOPE) HOFMANN LA ROCHE & CO AG F.
 XX
 PI Shao Z, Schmuck R, Kratzsch P, Kenkries J, Weisser H;
 XX
 DR WPI; 2003-383834/37.
 DR P-PSDB; ABR43480.
 XX
 PT New variant of an Erwinia-type creatinase modified relative to a wild-
 PT type creatinase having creatinase activity, useful for determining
 PT creatinase and/or creatine concentration in a sample.
 XX
 PS Example 4; Page 40-42; 51pp; English.
 XX
 CC The present invention describes a variant of an Erwinia-type creatinase
 CC (I) modified relative to a wild-type creatinase having creatinase
 CC activity. The variant comprises at least one amino acid substitution at a
 CC position of the fully defined 404 amino acid sequence given in ABR43467;
 CC these are selected from N130, M203, I278, I1304 and F395. Creatinase has
 CC the EC number EC 3.5.3.3 and is also known as creatine amidohydrolase.
 CC Also described is a reagent (II) for determining creatine comprising the
 CC Erwinia-type creatinase variant. The variant is useful for determining
 CC creatinine and/or creatine concentration in a sample. Measuring
 CC creatinine and/or creatine are useful for diagnosing uraemia, chronic
 CC nephritis, acute nephritis, gigantism, tonic muscular dystrophy and some
 CC other related diseases. The mutant enzymes have improved stability, lower
 CC conductivity and/or lower Km-values for creatine; they are much better
 CC suited to detection methods for creatine. The present sequence encodes a
 CC mutant Erwinia creatinase from the present invention
 CC
 SQ Sequence 1212 BP; 239 A; 391 C; 362 G; 220 T; 0 U; 0 Other;
 Query Match 80.7%; Score 978.4; DB 8; Length 1212;
 Best Local Similarity 88.0%; Pred. No. 8.2e-157;
 Matches 1066; Conservative 0; Mismatches 146; Indels 0; Gaps 0;

QY 481 TCGCTGGAAGAGCAGAAAGCTGATCCGCGAAGGCGCCCGCTGTGTGACGTGCGCGCGCG 540
 DB 481 TCGCTGGAAGAGCAGAAAGCTGATCCGCGAAGGCGCCCGCTGTGTGACGTGCGCGCGCG 540
 QY 541 GCTTCGCGCGCTGTCATCAAGGCGCGCGCTGCCGAGCATGAGTGGCGATTCGCACACCC 600
 DB 541 GCTTCGCGCGCTGTCATCAAGGCGCGCGCTGCCGAGCATGAGTGGCGATTCGCACACCC 600
 QY 601 AATGGAGATCCGCGAGATGCGCAATTCGTTCCCTTCGTGGAGCTGATGACCACTTGG 660
 DB 601 AATGGAGATCCGCGAGATGCGCAATTCGTTCCCTTCGTGGAGCTGATGACCACTTGG 660
 QY 661 ACCTGTTTCAGTCCGCGATCAACACGAGCGCGGCAACAATCCGCTCAACACCGCATC 720
 DB 661 ACCTGTTTCAGTCCGCGATCAACACGAGCGCGGCAACAATCCGCTCAACACCGCATC 720
 QY 721 GTGCAATCCGCGCATCTTTCGCTCAACACCTTCCGATGATCTTGGGCTTACTACCC 780
 DB 721 GTGCAATCCGCGCATCTTTCGCTCAACACCTTCCGATGATCTTGGGCTTACTACCC 780
 QY 781 GCGCTGAGCGCAGCTGTTGCGAGCATGTCATGTCAGCCAGCTCGACATCTGGAG 840
 DB 781 GCGCTGAGCGCAGCTGTTGCGAGCATGTCATGTCAGCCAGCTCGACATCTGGAG 840
 QY 841 AAGAACGTG3CCGTGCAATCGCGCGGCTCGAGCTGATCAACCGCGCGCGCTGCAAG 900
 DB 841 AAGAACGTG3CCGTGCAATCGCGCGGCTCGAGCTGATCAACCGCGCGCGCTGCAAG 900
 QY 901 GACATCGCCATGAGCTCAACGAGATGTACCGCGAGTGGAGACTGCTGAAGTACCGCTTC 960
 DB 901 GATATCGCCATGAGCTCAACGAGATGTACCGCGAGTGGAGACTGCTGAAGTACCGCTTC 960
 QY 961 TTGCGCTATGCGCACTCCTTGGCGGCTGTCGCTCACTACGCGCGAGCGCGCGCTG 1020
 DB 961 TTGCGCTATGCGCACTCCTTGGCGGCTGTCGCTCACTACGCGCGAGCGCGCGCTG 1020
 QY 1021 GAGCTGCGCGAGGACATGACACCGAGCTGAGAGCCCGCGCATGCTGTCTCATGAGCGCG 1080
 DB 1021 GAGCTGCGCGAGGACATGACACCGAGCTGAGAGCCCGCGCATGCTGTCTCATGAGCGCG 1080
 QY 1081 ATGTGATGCTGCGCGAGGCGCATGCGCGGCTGCGCGCGCTTATCGGAGACGACATCTTC 1140
 DB 1081 ATGTGATGCTGCGCGAGGCGCATGCGCGGCTGCGCGCGCTTATCGGAGACGACATCTTC 1140
 QY 1141 ATGCTGCGGAGGAGAGGTGCGGAGAACATCACCGGCTTCCGTTGGTCCGGAACAAC 1200
 DB 1141 ATGCTGCGGAGGAGAGGTGCGGAGAACATCACCGGATTCCTCCCTGCGGCTTGAACAAC 1200
 QY 1201 ATCATCCGCAAC 1212
 DB 1201 ATCATCCGCAAC 1212

RESULT 15
 ACC69520
 ID ACC69520 standard; DNA; 1212 BP.
 XX
 AC ACC69520;
 XX
 DT 21-JUL-2003 (first entry)
 XX
 DE Mutant Erwinia creatinase CTsd2 encoding DNA SEQ ID NO:19.
 XX
 KW Erwinia; creatinase; creatine amidohydrolase; enzyme; EC 3.5.3.3;
 KW Erwinia-type creatinase; creatine; creatinine; uraemia; gigantism;
 KW chronic nephritis; acute nephritis; tonic muscular dystrophy; mutant;
 KW gene; ds.
 XX
 OS Erwinia sp.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 7, 2005, 20:16:05 ; Search time 4424 Seconds
(without alignments)
10428.096 Million cell updates/sec

Title: US-10-807-228a-2

Perfect score: 1212

Sequence: 1 ATGACGACGACATCTTGA.....AACACACATCATCCGCAAC 1212

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:
1: gb_ests1:
2: gb_ests2:
3: gb_hrc:
4: gb_ests3:
5: gb_ests4:
6: gb_ests5:
7: gb_ests6:
8: gb_gsa1:
9: gb_gsa2:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	322.6	26.6	623	4	BI376242 BFLG3 000
2	221	18.2	470	1	AI906433 IL-BRT09-
3	207.8	17.1	494	6	CD307119 SctrPu691.
4	188.6	15.6	442	6	CD295681 SctrPu691.
5	182.6	15.1	821	6	CV214674 EST874384
6	135	11.1	670	7	CV214629 EST874339
7	122	10.1	519	6	CD296311 SctrPu691.
8	83.2	6.9	551	7	CP919042 BFL0531.
9	83.2	6.9	553	4	BI387857 BFL26 002
10	73.8	6.1	2529	9	CL964961 OsIFCC011
11	69.6	5.7	524	8	BZ894814 Hg4 0102
12	69.2	5.7	722	8	CB656525 OSJNEC10P
13	67.6	5.6	722	6	CB683938 OSJNEF12P
14	66	5.4	2538	9	CL963721 OsIFCC038
15	65	5.4	1368	9	CL948585 OsIFSB000
16	65	5.4	1368	9	CL948585 OsIFSB000
17	63.6	5.2	925	9	CL976880 OsIFCC043
18	61.6	5.1	1115	7	CK208301 FGAS02000
19	61.4	5.1	707	7	CF874011 trico35xc
20	61.4	5.1	732	7	CF886540 trico30xc
21	61.4	5.1	763	6	CB903555 trico35xc
22	61.4	5.1	802	6	CB902534 trico30xc
23	60.8	5.0	712	6	CB869465 AZ02.111L
24	60.8	5.0	812	6	CB659172 OSJNEC150

25	60.6	5.0	604	1	AU162766	AU162766
26	60.4	5.0	1575	9	CL979927	OsIFCC045
27	60.2	5.0	1116	9	CL963961	OsIFCC010
28	60.2	5.0	2022	9	CL973018	OsIFCC023
29	59.8	4.9	635	6	CA254345	SCBPR411
30	59.8	4.9	748	6	CA227320	SCULF301
31	59.8	4.9	1602	9	CL975927	OsIFCC043
32	59.6	4.9	993	9	CL977070	OsIFCC044
33	59.4	4.9	655	6	CA200834	SCRFPL107H
34	59	4.9	579	2	BE040798	OP1IG11 O
35	59	4.9	668	6	CD225047	CCCL1 37 H
36	59	4.9	690	2	BE041110	OF15H01 O
37	58.8	4.9	512	4	BM140351	WHE0474_a
38	58.8	4.9	614	2	BF202900	WHE1782 H
39	58.8	4.9	888	9	CG440030	OGVHP107H
40	58.6	4.8	719	7	CF487085	POL1 41 D
41	58.6	4.8	937	7	CG240162	OCYAE88TV
42	58.4	4.8	748	7	CN143339	WOUNDI 15
43	58.4	4.8	768	7	CN148454	WOUNDI 56
44	58.4	4.8	814	7	CN126847	RHOH1 19
45	58.2	4.8	1056	7	CK212443	FGAS02431

ALIGNMENTS

RESULT 1
BI376242
LOCUS
DEFINITION
BFLG3_000038 Amphioxus 5-6 hrs cDNA library (Name convention: BFLG or MPMGP498) Branchiostoma floridae cDNA clone MPMGP498A1518 5', mRNA sequence.

ACCESSION
BI376242
VERSION
BI376242.1 GI:30911206

KEYWORDS
SOURCE
ORGANISM
Branchiostoma floridae (Florida lancelet)
Branchiostoma floridae
Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;

REFERENCE
1 (bases 1 to 623)

AUTHORS
Panopoulou, G., Hennig, S., Groth, D., Krause, A., Poustecka, A., J., Herwig, R., Vingron, M., and Lehrach, H.

TITLE
New evidence for genome-wide duplications at the origin of vertebrates using an amphioxus gene set and completed animal genomes

JOURNAL
Genome Res. 13 (6A), 1056-1066 (2003)
MEDLINE
22683279
PUBMED
12799346

COMMENT
Contact: Panopoulou G
Laboratory 145, dept. Lehrach
Max-Planck-Institut fuer Molekulare Genetik
Inmestr.63-73, D-14195 Berlin, Germany
Tel: +49 30 8413 1235
Fax: +49 30 8413 1128
Email: panopoulou@molgen.mpg.de

The library was characterised by oligonucleotide fingerprinting (ONFP) to reduce sequencing redundancy. According to the ONFP procedure, clones giving the same hybridisation pattern with a battery of 200 Bmer oligonucleotides are grouped into clusters. One clone per cluster is selected for sequencing. The size of each cluster is an indicator of the frequency of a transcript in the analysed library. The cluster size as well the coordinates of the rest of the clones assigned to the same fingerprint cluster as the clone from which the above EST is generated is available at the amphioxus project site at: <http://www.molgen.mpg.de/amphioxus/>
Clones and filters are distributed via the Resource Center/Primary Database of the German Genome Project (<http://www.rzpg.de>)

PCR Primers
FORWARD: 5' CCCAGGCTTACATTATGCTCCGGCTG 3' (M1385P)
REVERSE: 5' GCTATTACGCGACGTGGGAAAGGGGATG 3' (M1385P)
Insert Length: 1200 Std Error: 0.00
Seq primer: 5'-CCGTCGGAATTCGGGGT-3' pSPORT3/86
High quality sequence stop: 623.

FEATURES	source	Location/Qualifiers
1..623		Location/Qualifiers
		/organism="Branchiostoma floridae"
		/mol_type="mRNA"
		/db_xref="taxon:7739"
		/clone="MPMP498A1518"
		/tissue_type="whole embryo"
		/dev_stage="5-6 hrs (gastrula stage)"
		/lab_host="E.coli, Xli blue"
		/clone_id="Amphioxus 5-6 hrs cDNA library (Name convention: BFG or MPMP498)"
		/note="Vector: pSPORT1; Site 1: SalI; XbaI, EcoRI (5'); Site 2: NotI, BamHI, HindIII (3'); Oligodt primed and directionally cloned in pSPORT vector using a NotI (5'-pGATTAATTCTTAAGTCGCGAAGCGGCCGCC (715-3' and a SalI 5'-TCGACCCACGGCGTCCG-3' adapters (Gibco BRL)."
ORIGIN		
Query Match	26.6%	Score 322.6; DB 4; Length 623;
Best Local Similarity	70.4%	Pred. No. 1.3e-60;
Matches 430;	Conservative 0;	Mismatches 181; Indels 0; Gaps 0;
QY	570	GCCGAGCATGAAAGTGCAGATCGCCACCAACCAATGCGATTCGCGAGATCGCCAAATC
DB	13	GCTGAGTACGAGGCTGGGAGTGCAGCGCTCCCAACATGATGATGGGAAATTGGCAGAC
QY	630	GTTCCCTCTTGTGAGCTGATGACACCTGACCTGTTTCCATGTCGGGATCAACACCGA
DB	73	CTACCCGATGGGGAATCATGAACACTTGGGCTGATTCAGTCAGGGATCAACATCGA
QY	690	CGGCGGCAAAATCCGATCAACCAACCGATGCTGATATCCGGGCAATCTTTGGCTCA
DB	133	CGGGGACACAACTCTGACGTCAAGCGCGGTGGCAGTGGGACATCTGTCACTAA
QY	750	CACCTTCCGATGATCTTGGCTACTACACCGCGCTGGAGGCGACGCTGTTCTGCGACA
DB	193	CACCTTGGCATATGATGGCGGGTACTACTTGGCATGAGAGGAACGATTCGTAACCA
QY	810	TGTCGATGACGCGACGCTCGACATCTGGAGAGAACTGGCCGTGATCGCCCGGACT
DB	253	CGCTCCGACGACACTCCGACTTTGGGAGATCACTGTAAGTATACGGCGGGGCT
QY	870	CGACTGATCAAGCCGGCGCGCGCTGGAAGACATTCGCCATTCAGCTCAACGAGATGA
DB	313	GGAAGTATTCGGCGGGAGGCGCGTGTGTGACATCCCACTGAGCTGAACGAGATGA
QY	930	CCGAGAGGAGACCTGTAAGTACCGCTCTTGGCTATGAGCCACTCTTCCGCGTCT
DB	373	CCGGGAAACGGGCTGCTGACATACCGCCCGATCGGCTTACCGGACATCTCTTCCGCGTAT
QY	990	GTCGCACTACTACGGTTCGAGAGCGCGAGTGGAGCTGCGAGGACATCGACCTGAGCT
DB	433	GTGACCTTACTACCGCAGAGGAAGCAGATTTGGAAGCTGAGAGGACGTCCATACGTTCT
QY	1050	GAAAGCCCGGCAATGTTGTCTCTCATGAGACCGATGTATGCTGCGGAGGGGATGCCCCG
DB	493	GGAAGCCCGGATGTCGTCTCATGAGACCGATGTATCAATTCGCCAGGCAATGCCCCG
QY	1110	TGCGGAGGCGCTATGCGAGGACGACATCTGATCGGGGAGGACGGTGGCGAGAACT
DB	553	AGCGGCGGTTATCCGGAGGACGACATCTTGTGTGACTGAGAGCGGGGCGAGAGATC
QY	1170	CACCGGCTTCC 1180
DB	613	TACANGGCTCC 623
RESULT 2		
LOCUS	AI906433	470 bp mRNA linear EST 30-MAR-2000
DEFINITION	II-BT109-280199-002 BT109 Homo sapiens CDNA, mRNA sequence.	
ACCESSION	AI906433	
VERSION	AI906433.1	GI:6496820

KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Bukaryotic; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 (bases 1 to 470) Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.P., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jorgensen,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE	Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE	20202863
PUBMED	10737800
COMMENT	Contact: Simpson A.J.G. Laboratory of Cancer Genetics Ludwig Institute for Cancer Research Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil Tel: +55-11-2704922 Fax: +55-11-2707001 Email: asimpson@ludwig.org.br This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/seq/gethtml.pl?fil=ILt2=IL-BT109-002.html&t=280199&t4=1) Seq primer: puc 18 forward. Location/Qualifiers
FEATURES	1..470 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /sex="female" /dev_stage="Adult" /clone_id="BT109" /note="Organ: Breast; Vector: puc18, Site_1: Sma1, Site_2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 199,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
ORIGIN	Query Match 18.2%; Score 221; DB 1; Length 470; Best Local Similarity 69.4%; Pred. No. 3, le-38; Matches 313; Conservative 0; Mismatches 137; Indels 1; Gaps 1;
	630 GTTCCCTCTGGAGCTGATGACACCTTGACCTGTTCACTCGGGCATCAACCGA 669 Db 455 GATCCCTGATGGCGGTGATTAACCTTGACATGGTTCACTCGGGTTCACACCGA 336 690 CGGCGGCGCAATTCGGGTACCAACCGCATCTCGGCATCCGGGACATCTTTTCGTCGA 749 395 TGGGGCCCATTAACCGGGGTACCGACCGCGGGTGGGAAAGGAATCTTCAGCTTCAA 336 750 CACCTTCGATGATCTTGGAG- TACTACACCGCGCTGAGAGCCACGCTTTCTGCGAC 808 335 TTGCTTCCCATGATCCCGGCATATTACACCGCGTTGAGCGCCACGCTTTCTCGACC 276 809 ATGTGATGACGCGCAGCTTGACATCTGGAGAGAAAGTGGCGGTGACATCGCCGGGCG 868 275 ACTGCCCGGACGAGTACTACGCTGTGGAGGACCAACGTCGAAGTGCAGAACCGGGGC 216 869 TCGAGCTGATCAAGCCGGGCGCGCTGCAAGACATCGGCATCGAGCTCAAGAAATCT 928 Db 215 TGAACCTGTGCGCCCGGCATCGCTGCGACGATATCGCTCGGGAATGAAAGAAATCT 156 929 ACGCGAGTGGGACCTGCAATGATACCGCTCTTGCGATATGCGCACTCTTCGGCGTGC 988

DB	155	TCCTGGGGCAGACACCTGGCTGCAGTACCGGACCTTGCGGTATGAGGCACATCGTTTGGCACT	96
QY	989	TGTCGCATCTACTACAGTGGTGGCCGAGCGCGCGTGTGAGCTGCGCGAGGACATCGACACCGAC	1048
DB	95	TGAGGCACCTACTATATGTGGCCGCGCNAAGGGGGGCTGTGAGACTGTGGCGAGCATCGACAGGTGC	36
QY	1049	TGAAGCCCGGCGATGTTGGTCTTCATCGAGCC	1079
DB	35	TGGAGCCGGGCGATGGTGGTGTCTATCAGGCC	5
RESULT 3			
LOCUS	CD307119/c		
DEFINITION	CD307119	494 bp	mRNA linear EST 16-SEP-2003
	Strp661.009486	Sea urchin larva cDNA library	MPMGp691
	Strongylocentrotus purpuratus	cDNA clone	
	MPMGp691B0990	MP1_SURUDI_90E9 5', mRNA sequence.	
ACCESSION	CD307119		
VERSION	CD307119.1	GI:34752168	
KEYWORDS	EST.		
SOURCE	Strongylocentrotus purpuratus		
ORGANISM	Strongylocentrotus purpuratus		
	Eukaryota; Metazoa; Echinodermata; Epleurozoa; Echinozoa;		
	Echinoidea; Euechinoidea; Echinacea; Echinoidea;		
	Strongylocentrotidae; Strongylocentrotus.		
REFERENCE	1 (bases 1 to 494)		
AUTHORS	Poustka,A.J., Groth,D., Hennis,S., Thamm,S., Cameron,A., Beck,A.,		
	Reinhardt,R., Herwig,R., Panopoulou,G. and Lehrach,H.		
TITLE	Generation, annotation, evolutionary analysis, and database		
	integration of 20,000 unique sea urchin EST clusters		
JOURNAL	Genome Res. 13 (12), 2736-2746 (2003)		
COMMENT	Contact: Poustka AJ		

FEATURES

mousekamo1gen.mpg.de
 was characterized by oligonucleotide fingerprinting
 to reduce sequencing redundancy. According to the ONF
 library, the clones that display the same hybridisation matrix with a
 set of 200 inner oligonucleotides are grouped into clusters. One
 ONF cluster is selected for sequencing. The size of each
 cluster is an indicator of the frequency of a transcript in the
 library. The cluster size as well as the coordinates of
 the clones assigned to the same ONF cluster as the clone from
 the above EST is generated and is available at the sea urchin
 web site at: <http://www.molgen.mpg.de/seq/seaurchin/>. cDNA
 and filters are distributed via the Resource Center/Primary
 of the German Human Genome Project (<http://www.rzpd.de>)

5'-seq
 5' CCCAAGGCTTACACTTTCATGCTCCGGCTCG 3' (M13RSP) 5'-seq
 5' GCGATTACGACGCTGGGAGAGAGGGGAGATGTG 3' (M13FSP) 3'-seq
 5' -CCGGTCCGGAATTCGCGGG-3' pSPort3/86
 5'-seq
 5'-seq
 Location/Qualifiers
 . . . 494

```

/organism="Strongyloides purpuratus"
/mol_type="mRNA"
/db_xref="taxon:7668"
/clone="MPMG691E0990.MPI.SUBUD1_90E9"
/tissue_type="whole larva"
/dev_stage="larva 2-3 weeks"
/lab_host="E.coli, Xli blue"
/notes="lib=Sea urchin larva cDNA library MPMG691"
/notes="lib=SPORT; sport; site 1: NotI; site 2: SalI; Random
primed and directionally cloned in SPORT vector using a
Sali 5'- TCACCTCAGCGCTCCG-3' adapter (5bco BRL)" and a

```

ORIGIN	
Query Match	17.1%; Score 207.8; DB 6; Length 494
Best Local Similarity	66.8%; Pred. No. 2.5e-35;

	Matches	312;	Conservative	0;	Mismatches	152;	Indels	3;	Gaps	1;
QY	730	GGCGACATCC	TTTGGCTGCACACACTT	TTCCCGATGATCTT	GGGCTACTACCGCGCTGGAG	789				
Db	476	GGTACGGC	ACATCGTCTTTAACGCTTCCCTCTTA	TGATACCGGCTATTTCT	CGGCGCTTGA	417				
QY	790	CGCAGCGTGAT	CTTCGACCAATGTGATGACGGCCAGC---	CTCGACATCT	GGAGAAAGAAC	846				
Db	416	CGAACCC	TTTTCTTGGAACACAGTGCCTTCCGACCGCATCT	TGAAGTGTGGCAAGTCAAC	357					
QY	847	GTGGCCGTGCAT	CGCCCGGGCTCTGAGCTGATCAAGCCGGCGCGCGCTG	CAAGACATC	906					
Db	356	TGCACAC	CTTACCGCCCGGGATGGA	ACTATCAAGCTGGTGTCA	ATGTGCGACGTC	297				
QY	907	GCATCGAG	CTCAACGAGATGTATACCGGAGGGGAACTG	CTGAGATACCGGCTCTT	CGGC	966				
Db	296	GCCTCGAG	TTTGAACGAGATGTATCCGGGAGAAAGAACTT	ACTACAGTACAGAAAGCTT	GGCC	237				
QY	967	TATGGC	CACTCTTGGCGGTGCTGTGC	CACTACTACG	GTGCGAGGCGCGGCTGGAGCTG	1026				
Db	236	TACGGG	CACTCTTGGGGTCTTATCCG	CACTACTACCGGCGCGAGAAAGAGCTT	CTTGAAGTGG	177				
QY	1027	CGGAGG	CAATTCGACACCGAGCTG	GAACCCGGCATGTGTGCTTCAT	TGAGCCGATG	1086				
Db	176	AGGGA	AAACATCGAGATCTGTATTATAC	AGCTGTGATGTGGTCTTTCAT	TGAGAACCAATCTC	117				
QY	1087	ATGTGTC	CGGAGGGCATACCCCGGTGCGCGCGCTAT	TGCGAGCAAGCAATCTG	TATGCTC	1146				
Db	116	ACGAT	TACAGACGGACGCTGTGGTG	CAGAGGCTTACAGGAGG	CAATGATATCATGTG	57				
QY	1147	GGGAG	AGACGATGCCGAGAAATCA	CGGAGCTTCCCGCTCGGTG	CCGGA	1193				
Db	56	ACTGA	AAAGGACCGCTGCTTCA	TATACCGGAATTTCTT	TACGAGACCGGA	10				

LOCUS	CD295681	442 bp	mRNA	linear	EST 16-SEP-2003
DEFINITION	<p> <i>Scirp</i>pe91.007076 Sea urchin larva cDNA library MPMGP691 <i>Strongylocentrotus purpuratus</i> cDNA clone MPMGP691E0990; MFL_SURUDI_90B3 3', mRNA sequence. </p>				
ACCESSION	CD295681				
VERSION	CD295681.1	GI:34746758			
KEYWORDS	EST.				
SOURCE	<i>Strongylocentrotus purpuratus</i>				
ORGANISM	<i>Strongylocentrotus purpuratus</i>				
	<p> Bakayote; Metazoa; Echinodermata; E]utherozoa; Echinozoa; Echinoidea; Euechinoidea; Echinacea; Echinoida; <i>Strongylocentrotidae</i>; <i>Strongylocentrotus</i>. 1 (bases 1 to 442) Poustka,A.J., Grotz,D., Hennig,S., Thamm,S., Cameron,A., Beck,A., Reinhardt,R., Herwig,R., Panopoulou,G. and Lehrach,H. Generation, annotation, evolutionary analysis, and database integration of 20,000 unique sea urchin EST clusters Genome Res. 13 (12), 2736-2746 (2003) Contact: Poustka AJ </p>				
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
COMMENT					

Laboratory 145, dept. Iernrach
Max-Planck-Institut fuer Molekulare Genetik
Inmestr. 63-73, D-14195 Berlin, Germany
Tel.: +49 30 8413 1235
Fax: +49 30 8413 1128
Email: pousetka@molgen.mpg.de
The library was characterised by oligonucleotide fingerprinting
(ONP) to reduce sequencing redundancy. According to the ONP
procedure, clones that display the same hybridisation matrix with a
battery of 200 8mer oligonucleotides are grouped into clusters. One
clone per ONF cluster is selected for sequencing. The size of each
cluster is an indicator of the frequency of a transcript in the
analysed library. The cluster size as well as the coordinates of
the other clones assigned to the same ONF cluster as the clone from
which the above EST is generated is available at the sea urchin
project web site at: http://www.molgen.mpg.de/ag_seaurchin/. cDNA
clones and filters are distributed via the Resource Center/Primary

Database of the German Human Genome Project (<http://www.rznp.de>)
 PCR Primers
 FORWARD: 5' CCCAGGCTTTACATTTATGCTCCGGCTCG 3' (M13RSP) 5'-seq
 BACKWARD: 5' GCTATTACGCGAGCGGAGGAGATG 3' (M13FSP) 3'-seq
 Seq primer: 5' GCTATTACGCGAGCGGAGGAGATG 3' (M13FSP)
 High quality sequence stop: 442.

FEATURES

source

1..442
 /organism="Strongylocentrotus purpuratus"
 /mol_type="mRNA"
 /db_xref="taxon:7668"
 /clone="MPMG691B0990.MPI SURUDI_90B9"
 /issue_type="whole larva"
 /dev_stage="larva 2-3 weeks"
 /lab_host="E.coli, XLI blue"
 /clone_lib="Sea urchin larva cDNA library MPMG691"
 /note="Vector: pSport1; Site 1: NotI; Site 2: SalI; Random
 primed and directionally cloned in pSport1 vector using a
 NotI (5'-pGACTAGTCTAGATCGGCGCGGCC (r)15-3' and a
 SalI 5'-TCGACCCACGCGTCCG-3' adaptors (Gibco BRL) "

ORIGIN

Query Match 15.6%; Score 188.6; DB 6; Length 442;
 Best Local Similarity 67.3%; Pred. No. 4.2e-31;
 Matches 282; Conservative 0; Mismatches 134; Indels 3; Gaps 1;

QY 730 GGGGACATCTTGGCTTCAACACCTTCCGATGATCTTGGGCTACTACACCGCGCTGGAG 789
 DB 20 GGTGACGCGATCGTCTTAACGCTTCCCTATATAGCGGCTATTTCTCGCGCTTGAA 79
 QY 790 CGGACCTTTTCGGACCAATGCGAAGCGCAGC---CTGACATCTGGGAGGAAG 846
 DB 80 CGAACCTTTTCTTGAACCAAGCTTCCGACCGCANTTTAAGTGGAGGTCAAC 139
 QY 847 GTGGCGGTGATCGCGCGGCTCGAGCTGATCAACCGCGCGCTGACAGACATC 906
 DB 140 TGGAGATTTCACCGCGCGGATGAGATCATCAAGCTTGCTCAATGCTCGACGTC 199
 QY 907 GGCATGAGCTCAACAGATGTAACGCGAGTGGAGCTTGTAAGTACCGCTCTTGGC 966
 DB 200 GGGCTGAGTGAACAGATGTACCGGAGAGAACTACTACATCAAGAGCTTGGC 259
 QY 967 TATGGCATCTCTTGGCGCGCTGCTGCTACTACTATGATGTCGAGACCGCGCTGAGAC 1026
 DB 260 TACGGGACATCTTCCGGGCTCTATGCACTACTACGCGCGAGAGAGTCTTGGAGTG 319
 QY 1027 CGGAGACATGCAACCGAGCTGAAGCCGCGATGATGATGCTCTCCATGAGCCGATGATG 1086
 DB 320 AGGAGAGATGCAAGCTGTATACAGCTGTATGCTGCTTCCATGAGCCACATCTC 379
 QY 1087 ATGCTCCGAGGAGCATGCGCGGCTGCGCGCTATGCGAGCAAGCATCTGATCT 1145
 DB 380 ACGATACCAAGACGAGCCTGCTGAGAGGCTACAGGAGCATATATATGAGTGT 438

RESULT 5

CV214674

821 bp mRNA linear EST 16-SEP-2004

LOCUS

CV214674

DEFINITION

EST874384 non-normalized T1 cDNA library Trichomonas vaginalis cDNA
 clone TYTC244 5' end, mRNA sequence.

ACCESSION

CV214674

VERSION

CV214674.1

GI:52161654

KEYWORDS

EST.

SOURCE

Trichomonas vaginalis

Trichomonas vaginalis

Eukaryota; Parabasalida; Trichomonadinae; Trichomonas.

ORGANISM

Trichomonadidae; Trichomonadinae; Trichomonas.

REFERENCES

1 (bases 1 to 821)

Carlton, J.M., Dyall, S., Johnson, P.J. and Fraser, C.M.
 The complete genome sequence of the sexually transmitted parasite
 Trichomonas vaginalis

JOURNAL

Unpublished (2004)

COMMENT

Contact: Jane Carlton

FEATURES

source

Parasite Genomics Group
 The Institute for Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-530-9319
 Fax: 301-838-0208
 Email: carlton@tigr.org
 Seq primer: lambda Triplex2.
 Location/Qualifiers
 1..821
 /organism="Trichomonas vaginalis"
 /mol_type="mRNA"
 /strain="T1"
 /db_xref="taxon:5722"
 /clone="TYTC244"
 /clone_lib="non-normalized T1 cDNA library"
 /note="Vector: Lambda Triplex2; Site 1: SfiI; Site 2:
 SfiI; T. vaginalis strain T1 library constructed from
 cDNA, made in lambda Triplex2. Inserts cloned
 unidirectionally in the SfiI and SfiI sites. Mass excision
 of library produced inserts in pTriplex2 plasmid. Inserts
 sequenced from both 5' and 3' ends using Triplex2
 sequencing primer and polydT 24 bp primer respectively."

ORIGIN

Query Match 15.1%; Score 182.6; DB 7; Length 821;
 Best Local Similarity 60.2%; Pred. No. 9.2e-30;
 Matches 327; Conservative 0; Mismatches 204; Indels 12; Gaps 1;

QY 35 ACAACGGCGAGAAAGATTATTCGCGTTTTCGATGCGAGATGACCGCGCCAAACG 94
 DB 279 ACAACGGCGAGAAAGATTATTCGCGCTTCTCAGACCGCGAGTTCGAGCTTCTCCG 338
 QY 95 ACGTTGCGCGCTGATGCGCAAGAACATGTCGATGCGCGCTGTACCTTATCACT 154
 DB 339 GCGTCCGCGCATCATGCGCGAAGACCTGACGCGCTCATCTCACCACTACACT 398
 QY 155 GCATCACTACTATTCCGCGCTGCTGATCTGATATTCGAGCAAGTACGCGATGATCA 214
 DB 399 CATCAAGTACTCTTCCGCTCTTCTTCACTTCTTGGCGGCTCTTACGCGATGATG 458
 QY 215 TCGACCAACAACGCAAGACGATTTTCGCGCGGATCGAGCGCGCGACCGCTGCGCC 274
 DB 459 TCACCAAGATGACACGCGTCACTATCAAGCAAGTATGACCGCGGATGCTTGGCGCC 518
 QY 275 GCACTTGGCGCAACATCACTACCTGATCGGATGCGCGCGCAATTCTATCGGCGC-- 332
 DB 519 GCACTTGGCGCAACATCGTGTACAGGATGCGCGCGGATGATCTATCACTACGCA 578
 QY 333 -----CGTGGCGCGAGTGAACCAAGCGCGCAAGCGCATGCGCATGATGATGACC 382
 DB 579 TCAGAGAGTCTTGGCGACCGCGGATATCAACCGCGCGCATTCGCGATGATGATC 638
 QY 383 ACGTCATCTGCACTTTCGCGCGCGCGAGCTGAGAAAGCTTACCGGCGCTGAGTTGTG 442
 DB 639 CGTGGCGCGAGTGAACCAAGATCAAGATCAAGCGCGCTTCTTCGAGAGGAGCGCTG 698
 QY 443 ACATCAAGCGCGCTTCAAGTATGATGATGCGGACATCAAGTGCCTCGAAGAGCAAGCTGA 502
 DB 699 ACGTTCGAGGACGAGATGCGCGAGCGATATCAAGTCCGCGGAGAGATGCGGCTCA 758
 QY 503 TCGGAGAAAGCGCGCGGTGTGACGTCGCGCGCGCGCTGCGCGGCTGCATCAAG 562
 DB 759 TCAAGACGATGCGCGCATTTGATCTTGGTGGTGGAGGCGATTCGCAAGCGCATCAAG 818
 QY 563 CCG 565
 DB 819 CCG 821

RESULT 6

CV214629

LOCUS

CV214629

DEFINITION

EST874339 non-normalized T1 cDNA library Trichomonas vaginalis cDNA
 clone TYTC244 5' end, mRNA sequence.

accession clone TVTC131 5' end, mRNA sequence.
 version CV214629
 keywors EST
 source Trichomonas vaginalis
 organism Trichomonas vaginalis
 reference Eukaryota; Parabasalida; Trichomonada; Trichomonadida;
 Trichomonadidae; Trichomonadinae; Trichomonas.
 authors 1 (bases 1 to 670)
 title Carlton, J.M., Dyall, S., Johnson, P.J. and Fraser, C.M.
 journal The complete genome sequence of the sexually transmitted parasite
 Trichomonas vaginalis
 comment Unpublished (2004)
 contact: Jane Carlton
 Parasite Genomics Group
 The Institute for Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850, USA
 tel: 301-530-9319
 fax: 301-838-0208
 email: carlton@tigr.org
 seq primer: lambda Triplex2.
 location/qualifiers
 1..670
 /organism="Trichomonas vaginalis"
 /mol_type="mRNA"
 /strain="T1"
 /db_xref="taxon:5722"
 /clone="TVTC131"
 /note="lib="non-normalized T1 cDNA library"
 /note="Vector: Lambda Triplex2; Site 1: SfiA; Site 2:
 SfiB; T. vaginalis strain T1 library constructed from
 cDNA, made in lambda Triplex2. Inserts cloned
 unidirectionally in the SfiA and SfiB sites. Mass excision
 of library produced inserts in pTriplex2 plasmid. Inserts
 sequenced from both 5' and 3' ends using Triplex2
 sequencing primer and polydT 24 bp primer respectively."

Query Match 11.1%; Score 135; DB 7; Length 670;
 Best Local Similarity 65.3%; Pred. No. 2.8e-19;
 Matches 198; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

QY 35 ACAACGGCGAGAAAGATTATTCGCGCTTTCCGATCCGAGATGACCCGCGCAAAAG 94
 DB 264 ACAACGGCGAGAAAGATTATTCGCGCTTTCCGATCCGAGATGACCCGCGCTTCG 323
 QY 95 AGCTTCGCGCGATGATGCGCAAGAACATGTGATGCGCGCTGTTACCTCTTATCACT 154
 DB 324 GCGTCGCGCGATGATGCGCAAGAACATGTGATGCGCGCTGTTACCTCTTATCACT 383
 QY 155 GCATCACTACTATTCGCGCTGCTGCTACTGCTATTTGCGAGCAAGTACGGCATGTCA 214
 DB 384 CCATCAAGACTACTGCGCTGCTGCTACTGCTATTTGCGCGCTCTCAAGCGCATGTG 443
 QY 215 TCGACCAACAACGCGCAGATTTGCGCGCGCATTCGAGCGCGCGCGCCTGCGCGC 274
 DB 444 TACACCAAGATGACGCGTCAATACGCAAAATTCACGCGCGGATGCGCTTGGCGCG 503
 QY 275 GAGGCTTCGCGCAACATCACTACCGACTGCGCGCGCAATTTCTATCGCGCGC 334
 DB 504 GAGGCTACGCGGAGAACTGTGTACAGGAGCTGCGCGGAGTAATCACTACACGCA 563
 QY 335 TGC 337
 DB 564 TCC 566

RESULT 7
 CD296311 519 bp mRNA linear EST 16-SEP-2003
 LOCUS CD296311
 DEFINITION Strongyloecentrotus purpuratus cDNA library MPMGp691
 MPMGp691N1215; MPI_SURUDI_15N12 5', mRNA sequence.

accession CD296311
 version CD296311.1 GI:34747388
 keywors EST
 source Strongyloecentrotus purpuratus
 organism Strongyloecentrotus purpuratus
 reference Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
 Echinoidae; Euechinoidae; Echinacea; Echinoida;
 Strongyloecentrotidae; Strongyloecentrotus.
 authors 1 (bases 1 to 519)
 title Pousetka, A.J., Groth, D., Hennig, S., Thamm, S., Cameron, A., Beck, A.,
 Reinhardt, R., Herwig, R., Panopoulou, G. and Lehnach, H.
 journal Generation, annotation, evolutionary analysis, and database
 integration of 20,000 unique sea urchin EST clusters
 Genome Res. 13 (12), 2736-2746 (2003)
 comment Contact: Pousetka, A.J.
 laboratory 145, dept. Lehnach
 Max-Planck-Institut fuer Molekulare Genetik
 Ihnestr. 63-73, D-14195 Berlin, Germany
 tel: +49 30 8413 1235
 fax: +49 30 8413 1128
 email: pousetka@molgen.mpg.de
 The library was characterised by oligonucleotide fingerprinting
 (ONP) to reduce sequencing redundancy. According to the ONP
 procedure, clones that display the same hybridisation matrix with a
 battery of 200 8mer oligonucleotides are grouped into clusters. One
 clone per ONP cluster is selected for sequencing. The size of each
 cluster is an indicator of the frequency of a transcript in the
 analysed library. The cluster size as well as the coordinates of
 the other clones assigned to the same ONP cluster as the clone from
 which the above EST is generated is available at the sea urchin
 project web site at: <http://www.molgen.mpg.de/sg/seaurchin/>. cDNA
 clones and filters are distributed via the Resource Center/Primary
 Database of the German Human Genome Project (<http://www.rzpd.de>)
 PCR Primers
 FORWARD: 5' CCCAGCGCTTACCTTATGCTTCGCGCTCG 3' (M13SP) 5'-seq
 BACKWARD: 5' GCTATTCAGCGCGATGCGGAGGAGATGCG 3' (M13PSP) 3'-seq
 Seq primer: 5'-CCGTCGCGGATTCGCCGCGT-3' pSport3/86
 High quality sequence stop: 519.

Query Match 10.1%; Score 122; DB 6; Length 519;
 Best Local Similarity 55.7%; Pred. No. 2e-16;
 Matches 274; Conservative 0; Mismatches 209; Indels 9; Gaps 2;

QY 10 GACATGTTGCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 69
 DB 24 GAGATGCAAAAGTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 83
 QY 70 GCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 129
 DB 84 GCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 143
 QY 130 GCGGCTGTTCACTCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 189
 DB 144 GCGTACTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 203
 QY 190 TTGCGACCAAGTACGATGATGATGATGATGATGATGATGATGATGATGATGAT 249
 DB 204 GTGCGTGTCCCTTACGATGATGATGATGATGATGATGATGATGATGATGATGAT 263

QY 250 ATCGACGGCGGCCAGCCCTGGCG-----CCGAGCTTGGCGGACAACTACTACCTACCC 303
 DB 264 GTTCAGACAGCGCCAGCGCTGGCGGAGATCAACCCGTGAGGACACATGTGGTCACTTACAG 323
 QY 304 GACTGGCGCGCGGACAAATTTCTATCGGCGGT---GCGCAGCTGACACCGGGCGGCAAG 360
 DB 324 GATTGGCATGGCACAACCTTCTGGAGCGCGTTCGGACCTGCTCGGCAACGCTTGAG 383
 QY 361 CGCATGGCATTCGATTCGACACGTCATCTGACTTCGCGCCGCGGAGCTCGAGAGCC 420
 DB 384 AGATGGAGCGGCAATTTGATCATATACCTCATGAGGAGAACTAAAGTATGAACAGCT 443
 QY 421 CTACCGGGCGGTGATGCTTCGACATCAGCCAGCCCTCGATGTGATGGCGACCATCAAG 480
 DB 444 GTATCTCGAGACGACCCGTGATGATGATACCATCATGATGAGTGAATCATCAAG 503
 QY 481 TGGCTCGAAGAG 492
 DB 504 TCCCCGAGAG 515

RESULT 8
 CFP919042 551 bp mRNA linear EST 05-NOV-2003
 LOCUS BFL0531.000127 Amphioxus 26 hrs CDNA library (Name convention: BFL26 or MPMGP531) Branchiostoma floridae cDNA clone
 DEFINITION MPMGP531.07115, BFL26_115L7 5', mRNA sequence.
 CFP919042
 CFP919042.1 GI:38190244
 EST.
 Branchiostoma floridae (Florida lancelet)
 Branchiostoma floridae
 Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae; Branchiostoma.
 1 (bases 1 to 551)
 Panopoulou, G., Hennig, S., Groth, D., Krause, A., Pousetka, A.J., Herwig, R., Vingron, M. and Lehrach, H.
 New evidence for genome-wide duplications at the origin of vertebrates using an amphioxus gene set and completed animal genomes
 Genome Res. 13 (6A), 1056-1066 (2003)
 JOURNAL 22683279
 MEDLINE 12799346
 PUBMED
 COMMENT
 CONTACT: Panopoulou G
 Laboratory 145, dept. Lehrach
 Max-Planck-Institut fuer Molekulare Genetik
 Ihnestr. 63-73, D-14195 Berlin, Germany
 Tel: +49 30 8413 1235
 Fax: +49 30 8413 1128
 Email: panopoulou@molgen.mpg.de
 The library was characterised by oligonucleotide fingerprinting (ONFP) to reduce sequencing redundancy. According to the ONFP procedure clones giving the same hybridisation pattern with a battery of 200 8mer oligonucleotides are grouped into clusters. One clone per cluster is selected for sequencing. The size of each cluster is an indicator of the frequency of a transcript in the analysed library. The cluster size as well as the coordinates of all clones assigned to the same fingerprint cluster are the clone from which the above EST is generated is available at the amphioxus project site at <http://www.molgen.mpg.de/amphioxus>.
 Clones and filters are distributed via the Resource Center/Primary Database of the German Genome Project (<http://www.rzpd.de>).
 PCR PRIMERs
 FORWARD: 5' CCCAGGCTTACACTTATGCTCCGGCTCG 3' (M13RSP)
 BACKWARD: 5' GCTATTAGCAGCGGCGGAGGAGATGTG 3' (M13FSP)
 Insert length: 1200 Std Error: 200.00
 Seq primer: 5'-CCGCTCCGGAATTCCTCCGGG-3' pSport3/86
 High quality sequence stop: 551.
 Location/Qualifiers
 1..551
 /organism="Branchiostoma floridae"
 /mol_type="mRNA"

ORIGIN
 Query Match 6.9%; Score 83.2; DB 7; Length 551;
 Best Local Similarity 57.8%; Pred. No. 7.1e-08;
 Matches 148; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

/strain="wild type"
 /db_xref="taxon:7739"
 /clone="MPMGP531.07115;BFL26_115L7"
 /issue_type="whole embryo"
 /dev_stage="26 hrs (neurula stage)"
 /lab_host="Escherichia coli, X1 blue"
 /clone_lib="Amphioxus 26 hrs cDNA library (Name convention: BFL26 or MPMGP531)"
 /note="Vector: pSport1; Site 1: SalI, KpnI, EcoRI (5'); Site 2: NotI, BamHI, HindIII (3'); OligodT primed and directionally cloned in pSport1 vector using a NotI (5'-TGACGAGTTCGATGCGGAGCGCGCC (T)15-3' and a SalI 5'-TGACCCACGCGCTCG-3' adapters (Gibco BRL)."

QY 26 TGAATGCGACACGGCGGAGAAATTTATGCGCGGTTTGGATGCGGATGACCCGCC 85
 DB 286 TGAATGCGATTAACGGCGGAGAAATTTATGCGCGGTTTGGATGCGGATGACCCGCC 145
 QY 86 GCCAAACGACGCTTGGCGGCTGATGCGCCAGAACATGTCGATGCGGCGCTTCACT 145
 DB 346 GGCTGACAGAGTGGGTGCTCATGCTCAGCAGTAATATGATGAGCTCTTCACTT 405
 QY 146 CTATATCATGCACTCACTATCTATTCGGGTGCTGATGCTATTTTCGACGCACTAG 205
 DB 406 CTATATCATGCACTCACTATCTATTCGGGTGCTGATGCTATTTTCGACGCACTAG 465
 QY 206 GATGATGATGACCAACAAACGACGAGATTTTCGCGGAGATGACGCGGCGGCGC 265
 DB 466 GGCTGCTGCTCACTGACGAAAGTCTCTTATCCCAACGCTTGAACGATGCTCAG 525
 QY 266 CCTGGCGCGGCGGCTT 281
 DB 526 CCTGGAGAGAGCAT 541

RESULT 9
 B1387857
 LOCUS BFL26 002560 Amphioxus 26hr CDNA library (Name convention: BFL26 or MPMGP531) Branchiostoma floridae cDNA clone MPMGP531.07115 5', mRNA sequence.
 DEFINITION
 B1387857
 B1387857.1 GI:30922696
 EST.
 Branchiostoma floridae (Florida lancelet)
 Branchiostoma floridae
 Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae; Branchiostoma.
 1 (bases 1 to 553)
 Panopoulou, G., Hennig, S., Groth, D., Krause, A., Pousetka, A.J., Herwig, R., Vingron, M. and Lehrach, H.
 New evidence for genome-wide duplications at the origin of vertebrates using an amphioxus gene set and completed animal genomes
 Genome Res. 13 (6A), 1056-1066 (2003)
 JOURNAL 22683279
 MEDLINE 12799346
 PUBMED
 COMMENT
 CONTACT: Panopoulou G
 Laboratory 145, dept. Lehrach
 Max-Planck-Institut fuer Molekulare Genetik
 Ihnestr. 63-73, D-14195 Berlin, Germany
 Tel: +49 30 8413 1235
 Fax: +49 30 8413 1128
 Email: panopoulou@molgen.mpg.de
 The library was characterised by oligonucleotide fingerprinting (ONFP) to reduce sequencing redundancy. According to the ONFP procedure, clones giving the same hybridisation pattern with a battery of 200 8mer oligonucleotides are grouped into clusters. One

QY 652 GACACCTGACCTGTGTTCCAGTGGGATCAACACGACGCGGACCAATCCGGTACC 711
 DB 256 GTGGGCTCGGCGGCGGACGCTGACAGAACCGCCAGGTCTTCTTCCCGACCTCC 315
 QY 712 AACCGATCGTGCAATCCGCGCATCTCTTTCGCTCAACCTTCCGATGATCTTGGC 771
 DB 316 GTGACCTCGTGGGTGGCCAGCGCGTGCCTCCGCTCCGCGCATCAACGAGCTACGGC 375
 QY 772 TACTACACCGCGCTGAGGCGACGCTTCTTGGGACCATGTCCATGACGCGCTTGAC 831
 DB 376 CTGACGCGGCTGACGCTGACCTCAAGACCTTCAACACGACGCGCGCGCTGAC 435
 QY 832 ATCTGGAGAGAAAGCTGGCGCTGACATCGCGGCGGCTGAGCTGATCAAGCGGGCGC 891
 DB 436 ACCTGTGAGAGTGCATCGCGCGCTCTCTTCAACGAGCTCAAGGCGCGGCAACCGAATC 495
 QY 892 CGCTGCAAGACATCGCATCGAGCTCAACGAGATGACCGGAGTGGGACTGTGTAAG 951
 DB 496 ACCACTCCATCGCGCGCTTCAAGAGACGCGTGGTGAAGCGCTACTACAGCGCGTGG 555
 QY 952 TACCGCTCTTGGGCTATGCGCACTCTTGGCGGTGTGCTCACTACTACGCTCGGAG 1011
 DB 556 CGCGGCTACGCGCGGCTGATGACCTCATCACTTCAAGCTTCAAGCTACGCGCAAC 615
 QY 1012 GCCGGGTGAGGCTGGCGGACGATGACACGAGCTGAAGCGCGCATGTTGCTTC 1071
 DB 616 ACCGACGTGCGGACGATGATGATGTTCTTACGACGAGAGCGCGGAATACCGGCGGC 675
 QY 1072 ATGAGCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1109
 DB 676 AAGGTCTCACGAGCTTCAAGACGCGGAGCTGCGCG 713

RESULT 13
 LOCUS CB683938 722 bp mRNA linear EST 09-APR-2003
 DEFINITION OSJUNEf12P21.f OSJUNEf Oryza sativa (japonica cultivar-group) cDNA
 clone OSJUNEf12P21 5', mRNA sequence.

ACCESSION CB683938
 VERSION CB683938
 KEYWORDS EST.
 SOURCE Oryza sativa (japonica cultivar-group)
 ORGANISM Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE 1 (bases 1 to 722)
 Jantasuriyarat, C., Lu, G., Gowda, M., Hatfield, J., Zhou, B., Mazur, E.,
 Kudrna, D., Dean, R., Soderlund, C., Wing, R. and Wang, G.
 Large-scale identification of ESTs involved in the interaction
 between rice and Magnaporthe grisea
 Unpublished (2003)

JOURNAL COMMENT
 TITLE Arizona Genomics Institute
 CONTACT: Rod Wing
 UNIVERSITY OF ARIZONA
 BIOLOGICAL SCIENCES WEST, 448A, P.O. Box 210088, Tucson, AZ
 85721-0088, USA
 TEL: 520 626 3967
 FAX: 520 621 9288
 EMAIL: http://genome.arizona.edu
 PCR PRIMER
 FORWARD: gta aac cga cgg cca gtc
 BACKWARD: gga aac agc tat gac cat g
 PLACE: 12 row: P column: 21
 Seq primer: gta aac cga cgg cca gtc.

FEATURES
 source
 1..722
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="OSJUNEf12P21"

/cissue_type="Leaf"
 /dev_stage="3 week"
 /lab_host="DH10B"
 /clone_idb="OSJUNEf"
 /note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
 XhoI; Uninfected Control"

Query Match 5 6%; Score 67.6; DB 6; Length 722;
 Best Local Similarity 44.8%; Pred. No. 0.0002;
 Matches 259; Conservative 0; Mismatches 319; Indels 0; Gaps 0;

QY 532 GGGGCGCGGCGCTGCGCGGCTGCGCATCAAGCGCGGCTCCGAGCATGATGAGCATC 591
 DB 142 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 201
 QY 592 GCCACCAACATGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 651
 DB 202 AACCTGTCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 261
 QY 652 GACACCTGACCTGTGTTCCAGTGGGACATCAACGAGCGGCGGCGGCGGCGGCGGCGG 711
 DB 262 GTGGGCTTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 321
 QY 712 AACCGATCGTGCAATCCGCGGACATCTTTCGCTCAACCTTCCGATGATCTTGGC 771
 DB 322 GTGACCTCGTGGGTGGCCAGCCCGCTCTCCGCTCCGCTCCGCTCCGCTCCGCTCCG 381
 QY 772 TACTACACCGCGCTGAGGCGACGCTTCTGCGACATGTGATGATGATGATGATGATG 831
 DB 382 CTGACGCGGCTGACGCTGACATCAAGACCTTCAACGAGCGGCGGCGGCGGCGGCGG 441
 QY 832 ATCTGGAGAGAAAGCTGGCGCTGATGATGATGATGATGATGATGATGATGATGATG 891
 DB 442 ACCTGTGAGAGTGCATCGCGCGCTCTCTTCAACGAGCTCAAGCGCGGCGGCGGCGG 501
 QY 892 CGCTGCAAGACATCGCATCGAGCTCAACGAGATGATGATGATGATGATGATGATGATG 951
 DB 502 ACCACTCCATCGCGCGCTTCAAGAGACGCGGTGTCAGCGCTACTACAGCGCGTGG 561
 QY 952 TACCGCTCTTGGGCTATGCGCACTCTTGGCGGTGTGCTCACTACTACGCTCGGAG 1011
 DB 562 CGCGGCTACGCGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 621
 QY 1012 GCCGGGTGAGACTGCGGCGGACATGACACCGAGCTGAAGCGCGCATGTTGCTTC 1071
 DB 622 ACCGACGTGCGGACGATGATGATGATGATGATGATGATGATGATGATGATGATG 681
 QY 1072 ATGAGCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1109
 DB 682 AAGGTCTCACGAGCTTCAAGACGCGGAGCTGCGCG 719

RESULT 14
 LOCUS CL963721 2538 bp DNA linear GSS 21-SEP-2004
 DEFINITION OIRFC038554 Oryza sativa Express Library Oryza sativa (indica
 cultivar-group) genomic, genomic survey sequence.

ACCESSION CL963721
 VERSION CL963721
 KEYWORDS GSS.
 SOURCE Oryza sativa (indica cultivar-group)
 ORGANISM Oryza sativa (indica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE 1 (bases 1 to 2538)

JOURNAL AUTHORS Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M.,
 Zhao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L.,
 Wong, G. K. S., Deng, X. W. and Wang, J.
 TITLE An analysis of transcriptional regulation of the rice genome and
 JOURNAL Unpublished (2004)

COMMENT Contact: Chen Chen
Department of Bioinformatic
Beijing Institute of Genomics
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-80481576
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Class: exon-trapped.

FEATURES
source
1..2538
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="genomic DNA"
/db_xref="taxon:39946"
/clone_lib="Oryza sativa Expressed Library"
/note="Oryza sativa exon trapped genomic sequences"

ORIGIN

Query Match 5.4%; Score 66; DB 9; Length 2538;
Best Local Similarity 44.8%; Pred. No. 0.00048;
Matches 295; Conservative 0; Mismatches 360; Indels 3; Gaps 1;

Oy 408 GCTCGAGGAAGCCCTACCGGCGCTCGAGTTCTGTCATCAGCCAGCCCTTCATGTGAT 467
Db 1167 GCGCGTGAACGCGCTCTCTCAGCGCGCTGAGGTCTCAAGATGAGCACTCGGTGGGAG 1226
Oy 468 GCGCACATCAAGTGTCTGTAAGAGCAAGCTGATCCGGAAGCGCGCGGTGTGTA 527
Db 1227 CTTCAACGCGAGTTCTGCGTGGACGGAAGAGCGCGACGCGGAGCGCGCA 1286
Oy 528 CCGCGCGCGCGCGCGCTGCGCGCTGCCATCAAGCGCGCGTCCGAGCATGAAGTGC 587
Db 1287 GCGCGTGGGCGGCGGCTGCGAGTATGTTCCGCGCGTTCGCGCGCGTGGAGCCAT 1346
Oy 588 GATCGCACCAATCATGATGATATCGCGAGATCGCAATGCTTCCCTTCGTGAGCT 647
Db 1347 GCGCGTCAAGTGTACAGCGCGCGCGAGATCGGAGCGCGGAAAGCTTCTCGTGTG 1406
Oy 648 GATGACACCTGAGCTGTTCTCAATCGCGCATCAACCGAGCGCGCGCATTCGGT 707
Db 1407 GCTGCTCCCATTCACACCGCGCAATCTTTCACACAGAGCGCGGTTCAGCAAGAG 1466
Oy 708 CACCAACCGCATTCGTCATTCGCGCGCATTCCTTTCGTCACACCTTCGATGTATCT 767
Db 1467 GCGCTACACCTTCTCTCCACCTTCGCGCTTCGCGCTTCTTCTTCGCGAGATCA 1526
Oy 768 CGGCTACTACCGCGCTGAGCGACGCTGTTCTGCAACATGTGATGAGCGCCAGCT 827
Db 1527 GCGCGCGACCAAGAACTTGAAGAGAGCGCATCATCGCGCGCTTCGCGCAAGT 1586
Oy 828 CGACATCTGGAGAGAAAGACGTGCGCTGATTCGCGCG--GCTGAGCTGATCAAGCC 884
Db 1587 CTACATCGCGAGATGACGAGCGCACCAAGTTCGCGTGAAGCGGGAACCCGAGTTC 1646
Oy 885 GCGCGCGCGCTGAGAGATCGCCATTCGAGCTCAACGATTCACCGGAGTGGAGCT 944
Db 1647 GGAGCAAGGAGATCAAGAGTTCACACGAGATTCAGATCTGTCAAGCTCCGCCACCG 1706
Oy 945 GCTGAAGTACCGCTCTTGGCTATGAGCACTCTTCGCGTGTGCTGCACTACTACGG 1004
Db 1707 CCACTCGTCTCTCTCATGCGCTACGCGACGAGAACGCGAGATATCTCTGTCTACGA 1766
Oy 1005 TCGCGAGCGCGCGTGAAGCTGCGCGAGGACATGACCGAGCTGAAGCGCGCATG 1062
Db 1767 GTACATGACCAATGCGCGCTTCGCGACCACTACGCGAGAGACCTCGCGCTTG 1824

RESULT 15
CL948585 1368 bp DNA linear GSS 21-SEP-2004
LOCUS OSIRSB000026 Oryza sativa Expressed Library Oryza sativa (indica
DEFINITION cultivar-group) genomic, genomic survey sequence.
ACCESSION CL948585

VERSION CL948585.1 GI:52360594
KEYWORDS GSS
SOURCE Oryza sativa (indica cultivar-group)
ORGANISM Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE
1 (bases 1 to 1368)
Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M.,
Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L.,
Wong, G. K. S., Deng, X. W. and Wang, J.
An analysis of transcriptional regulation of the rice genome and
its comparison to Arabidopsis
Unpublished (2004)

JOURNAL
CONTACT Contact: Chen Chen
Department of Bioinformatic
Beijing Institute of Genomics
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-80481576
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Class: exon-trapped.

FEATURES
source
1..1368
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="genomic DNA"
/db_xref="taxon:39946"
/clone_lib="Oryza sativa Expressed Library"
/note="Oryza sativa exon trapped genomic sequences"

ORIGIN

Query Match 5.4%; Score 65; DB 9; Length 1368;
Best Local Similarity 44.4%; Pred. No. 0.00077;
Matches 263; Conservative 0; Mismatches 330; Indels 0; Gaps 0;

Oy 587 CGATGCCACCAACCATGATGATCGCGAGATCGCAATGTTCCCTTCGTGAGC 646
Db 197 CGAGACCGCCACGAGAGCGAGATCCCGCTTCACTCGCGAGATCGTCAAGCTTC 256
Oy 647 TGATGACACCTGACCTGTTCTCAATCGCGCATCAACGAGCGCGCATTCGG 706
Db 257 TCGGAGACTCACCGCGAGAGCTAGCCACAGCGCGCGCTTCGCGAGAGCGAG 316
Oy 707 TCACCAACCGCATTCGTCATTCGCGCGCATTCCTTTCGTCACACCTTCGAGTATCT 766
Db 317 GTACCAACCGCATTCGTCATTCGCGCGCATTCCTTTCGTCACACCTTCGAGTATCT 376
Oy 767 TCGGCTACTACCGCGCTGAGCGACGCTGTTCTGCGACCATGTGATGAGCGCCAGCT 826
Db 377 GCGCGCGCTACCGCGCGGTGCTTCGCGCGCGCGCGCGCTTCGTCAGCGCGCTCGG 436
Oy 827 TCGACATCTGGAGAGAAAGCTGCGCTGATTCGCGCGCGCTTCGAGCTGATCAAGCCG 886
Db 437 ACATGCGCATCAATGCTGCGCGCGCATGACACGCGCTTCGCGTCAAGCGCACCGGCT 496
Oy 887 GCGCGCGCTGAGAGATCGCCATTCGAGCTCAACGATTCACCGGAGTGGAGCT 946
Db 497 TGTGTACGTCACCAACATGCTGCTGCGCATCAAGAGTCTCTCGCACCTTCAGCGCG 556
Oy 947 TGAAGTACCGCTCTTGGCTATGAGCACTCTTCGCGTGTGCTGCACTACTACGG 1006
Db 557 TATCTGTGTGACATTCAGCGCGCACACGCGGAGCGCGTGAAGATGCTTCTTGAAGT 616
Oy 1007 GCGAGCGCGCGTGAAGCTGCGCGAGGACATGACCGAGCTGAAGCTCGCGCATG 1066
Db 617 CCAACCGGCTCATGACGCTGTGTTCAACCGGTACGCGCAAAATCAACGCGCCCAACAACT 676
Oy 1067 TCTTCATGAGAGCCATGATGATGCTGCGGAGAGGATGCGGAGCGCGGCTATCGCG 1126
Db 677 TCTTCGCGCGCTTCGCGCGCTTCGCGCAAGATTCGCGCGCGCGCAAGCATCAAGCG 736
Oy 1127 ACACGACATCTGATGCTGCGGAGAGACGCTGCGAGAACATCAACGCGCTTC 1179

Db 737 TGAACGTGCCGCTGSAAGCCGCGCTCCGCGACGACGTGTACCAACACGCTGTTTC 789

Search completed: July 7, 2005, 23:30:21
Job time : 4430 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 7, 2005, 20:27:44 ; Search time 236 Seconds
(without alignments)
8403.252 Million cell updates/sec

Title: US-10-807-228a-2

Perfect score: 1212
Sequence: 1 ATGACTGACGACATGTTGCA.....AACACACATCATCGGCAC 1212

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents NA:*

1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
4: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
5: /cgn2_6/prodata/1/ina/5C_COMB.seq:*
6: /cgn2_6/prodata/1/ina/5D_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1210.4	99.9	1212	1	US-09-940-941-2
2	1210.4	99.9	1212	2	US-08-799-897-2
3	1207.2	99.6	1215	2	US-08-947-7268-1
4	74.8	6.2	858	4	US-09-252-991A-4171
5	74.8	6.2	963	4	US-09-252-991A-4435
6	74.8	6.2	1176	4	US-09-252-991A-4314
7	64	5.6	3157	2	US-08-939-002A-1
8	64	5.3	1294	3	US-09-025-691-2
9	63.8	5.3	18686	4	US-09-902-540-1206
10	62	5.1	2064	1	US-08-343-428-1
11	60.6	5.0	1599	4	US-09-475-515-54
12	60.6	5.0	2112	4	US-09-475-515-56
13	60.6	5.0	2112	4	US-09-475-515-57
14	60.6	5.0	2181	4	US-09-475-515-58
15	60.6	5.0	2634	4	US-09-475-515-64
16	60.6	5.0	4284	4	US-09-902-540-3289
17	60.6	5.0	4766	4	US-09-475-515-73
18	60.6	5.0	17727	4	US-09-902-540-1152
19	60	5.0	1822	4	US-09-949-016-1464
20	60	5.0	1965	4	US-09-220-132-21
21	60	5.0	9399	4	US-09-949-016-13206
22	59.2	4.9	1056	4	US-09-266-965-66
23	59.2	4.9	1293	2	US-08-924-440-1
24	59.2	4.9	53500	4	US-09-266-965-76
25	58	4.8	2712	3	US-09-025-691-4
26	57.8	4.8	3780	4	US-09-902-540-9287
27	57.8	4.8	9873	4	US-09-902-540-9932

28	57.4	4.7	1434	4	US-09-902-540-4622	Sequence 4622, Ap
29	57.4	4.7	6776	4	US-09-902-540-884	Sequence 884, Ap
30	57.4	4.7	14809	4	US-09-902-540-1032	Sequence 1032, Ap
31	57.2	4.7	1695	4	US-09-902-540-4319	Sequence 4319, Ap
32	57.2	4.7	2277	4	US-09-902-540-9434	Sequence 9434, Ap
33	57.2	4.7	11358	4	US-09-902-540-1075	Sequence 1075, Ap
34	57.2	4.7	12730	4	US-09-902-540-1029	Sequence 1029, Ap
35	56.6	4.7	1476	3	US-09-434-288-12	Sequence 12, Ap
36	56.4	4.7	876	4	US-09-902-540-5288	Sequence 5288, Ap
37	56.4	4.7	1005	4	US-09-252-991A-1549	Sequence 1549, Ap
38	56.4	4.7	1590	4	US-09-252-991A-1617	Sequence 1617, Ap
39	56.4	4.7	1659	4	US-09-252-991A-1566	Sequence 1566, Ap
40	56.4	4.7	30783	4	US-09-902-540-1258	Sequence 1258, Ap
41	55.8	4.6	1644	4	US-09-902-540-9071	Sequence 9071, Ap
42	55.8	4.6	7719	4	US-09-902-540-969	Sequence 969, Ap
43	55.4	4.6	651	4	US-09-902-540-4369	Sequence 4369, Ap
44	55.4	4.6	2655	1	US-08-471-033-17	Sequence 17, Ap
45	55.4	4.6	2655	1	US-08-471-033-26	Sequence 26, Ap

ALIGNMENTS

RESULT 1
US-09-940-941-2
Sequence 2, Application US/0940941
Patent No. RE38687
GENERAL INFORMATION:
APPLICANT: Sogabe, Atsushi
Hattori, Takashi
Nishiyama, Yoshiaki
Kawamura, Yoshiaki
TITLE OF INVENTION: NOVEL CREATIVE AMIDNOHYDROLASE, PRODUCTION
THEREOF AND USE THEREOF
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: LEYDIG, VOIT & MAYER, LTD.
STREET: Two Prudential Plaza, Suite 4900
CITY: Chicago
STATE: Illinois
COUNTRY: US
ZIP: 60601-6780
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/940,941
FILING DATE: 28-Aug-2001
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/799,897
FILING DATE: 13-FEB-1997
APPLICATION NUMBER: JP 25435/1996
FILING DATE: 13-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: Robert F. Green
REGISTRATION NUMBER: 27555
REFERENCE/DOCKET NUMBER: 78064
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 616-5600
TELEFAX: (312) 616-5700
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1212 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA
ORIGINAL SOURCE:
ORGANISM: Alcaligenes faecalis

STRAIN: TE3581 (FERM P-14237)
FEATURE:
NAME/KEY: CDS
LOCATION: 1 to 1212
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-940-941-2

Query Match 99.9%; Score 1210.4; DB 1; Length 1212;
Best Local Similarity 99.9%; Pred. No. 1.8e-242;
Matches 1211; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGACTGACGACATGTTGACGATGTAATGCAACGCGGAGAAATATTTCGCGG 60
DB 1 ATGACTGACGACATGTTGACGATGTAATGCAACGCGGAGAAATATTTCGCGG 60
QY 61 TTTTGGATGCGAGATGACCCGCGCCAAAAGAGTTGCGGCTGATGGCCAAAG 120
DB 61 TTTTGGATGCGAGATGACCCGCGCCAAAAGAGTTGCGGCTGATGGCCAAAG 120
QY 121 AATGTCGATGCGGCGCTGTTACCTTTATCACTGATCACTACTATTTCGCGCTG 180
DB 121 AATGTCGATGCGGCGCTGTTACCTTTATCACTGATCACTACTATTTCGCGCTG 180
QY 181 TACTGCTATTTCGAGCGCAAGTACGCGATGTCATGACCAACAGCGCAAGAT 240
DB 181 TACTGCTATTTCGAGCGCAAGTACGCGATGTCATGACCAACAGCGCAAGAT 240
QY 241 TCGGCGCGGATGACGCGCGCGCGAGCTTCGCGCGGCAACATCACTAC 300
DB 241 TCGGCGCGGATGACGCGCGCGCGAGCTTCGCGCGGCAACATCACTAC 300
QY 301 ACCGACTGCGCGCGGCAATTTCTATCGCGCGCTGCGCACTGACCGCGCGCAAG 360
DB 301 ACCGACTGCGCGCGGCAATTTCTATCGCGCGCTGCGCACTGACCGCGCGCAAG 360
QY 361 CGCATCGGATGAGTTGACACACGCACTGCACTTCGCGCGGCACTGAGAGG 420
DB 361 CGCATCGGATGAGTTGACACACGCACTGCACTTCGCGCGGCACTGAGAGG 420
QY 421 CTACCGGCGGATGAGTTGCTGACATCAACGCGCGCTGATGATGCGCACTCAAG 480
DB 421 CTACCGGCGGATGAGTTGCTGACATCAACGCGCGCTGATGATGCGCACTCAAG 480
QY 481 TCGCTGAAAGACAGAGCTGATTCGCGGAGGCGCGCTGCTGAGCTCGCGCGCG 540
DB 481 TCGCTGAAAGACAGAGCTGATTCGCGGAGGCGCGCTGCTGAGCTCGCGCGCG 540
QY 541 GCGTCGCGGCGGATCAAGGCGCGCGCGCGGATGAGTGGCGATCGCAACGCG 600
DB 541 GCGTCGCGGCGGATCAAGGCGCGCGCGCGGATGAGTGGCGATCGCAACGCG 600
QY 601 AATGCGATGATCCGCGAGATCGCAATCGTTCCCTTCGTTGAGTGAAGCACTTGG 660
DB 601 AATGCGATGATCCGCGAGATCGCAATCGTTCCCTTCGTTGAGTGAAGCACTTGG 660
QY 661 ACTGCTTCAGTCCGCGCATCAACCGCGCGCGCAATCGCTGCAACCGCGCAT 720
DB 661 ACTGCTTCAGTCCGCGCATCAACCGCGCGCGCAATCGCTGCAACCGCGCAT 720
QY 721 GCGCAATCCGCGGACATCTTTTCGCGCAACCTTCGCGATATCTTGGGCTACTAC 780
DB 721 GCGCAATCCGCGGACATCTTTTCGCGCAACCTTTCCGATGATCTTGGGCTACTAC 780
QY 781 GCGCTGAGAGCGACGCTGTTTCGCGACATGTCGATGACGCGACCTCGACATCTGGAG 840
DB 781 GCGCTGAGAGCGACGCTGTTTCGCGACATGTCGATGACGCGACCTCGACATCTGGAG 840
QY 841 AAGAACTGCGCGTGCATGCGCGCGGCTGAGCTGATCAAGCGCGCGCGCTGCAAG 900
DB 841 AAGAACTGCGCGTGCATGCGCGCGGCTGAGCTGATCAAGCGCGCGCGCTGCAAG 900
QY 901 GACATCGCATGAGCTCAACGAGATGTAACGCGAGTGGAGCTGCTGAATCGGCTCC 960
DB 901 GACATCGCATGAGCTCAACGAGATGTAACGCGAGTGGAGCTGCTGAATCGGCTCC 960

DB 901 GACATCGCATGAGCTCAACGAGATGTAACGCGAGTGGAGCTGCTGAATCGGCTCC 960
QY 961 TTCGCTATGCGCACTCTTTGCGGCTGCTGTCCTACTATCACTGCGGAGCGCGCTG 1020
DB 961 TTCGCTATGCGCACTCTTTGCGGCTGCTGTCCTACTATCACTGCGGAGCGCGCTG 1020
QY 1021 GAGCTGCGGAGGACATGACACCGAGCTGAAGCGCGGATGATGCTCTCCATGAGCGG 1080
DB 1021 GAGCTGCGGAGGACATGACACCGAGCTGAAGCGCGGATGATGCTCTCCATGAGCGG 1080
QY 1081 ATGTCGATGCTGCGGAGGAGATGCGCGGCTGCTATTCGAGACAGACATCTCTG 1140
DB 1081 ATGTCGATGCTGCGGAGGAGATGCGCGGCTGCTATTCGAGACAGACATCTCTG 1140
QY 1141 ATTCGTCGCGGAGGACGCTGCGGAGACATCAACGCGCTTCGCTGCTGCGGAAACAAC 1200
DB 1141 ATTCGTCGCGGAGGACGCTGCGGAGACATCAACGCGCTTCGCTGCTGCGGAAACAAC 1200
QY 1201 ATCATCCGCAAC 1212
DB 1201 ATCATCCGCAAC 1212

RESULT 2
US-08-799-897-2
Sequence 2, Application US/08799897
Patent No. 6080553
GENERAL INFORMATION:
APPLICANT: Sogabe, Atsushi
APPLICANT: Hattori, Takashi
APPLICANT: Nishiyama, Yoshihisa
APPLICANT: Kawamura, Yoshihisa
TITLE OF INVENTION: NOVEL CREATINE AMIDINOHYDROLASE, PRODUCTION
NUMBER OF SEQUENCES: 3
TITLE OF INVENTION: THERBOF AND USE THEREOF
CORRESPONDENCE ADDRESS:
ADDRESSEE: LEYDID, VOIT & MAYER, LTD.
STREET: Two Prudential Plaza, Suite 4900
CITY: Chicago
STATE: Illinois
COUNTRY: US
ZIP: 60601-6780
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/799,897
FILING DATE: 13-FEB-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 25435/1996
FILING DATE: 13-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: Robert F. Green
REGISTRATION NUMBER: 27555
REFERENCE/DOCKET NUMBER: 78064
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 616-5600
TELEFAX: (312) 616-5700
TELEX: 25-3533
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1212 base pairs
TYPE: nucleic acid
STRANDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA
ORIGINAL SOURCE:
ORGANISM: Alcaligenes faecalis
STRAIN: TE3581 (FERM P-14237)
FEATURE:

NAME/KEY: CDS
LOCATION: 1 to 1212
US-08-799-897-2

Query Match 99.9%; Score 1210.4; DB 3; Length 1212;
Best Local Similarity 99.9%; Pred. No. 1.8e-242;
Matches 1211; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ATGACGAGAGCAATGTTGCAACGATGAATGGCAACAGCGGAGAAATTTTGGCCG 60
1 ATGACGAGAGCAATGTTGCAACGATGAATGGCAACAGCGGAGAAATTTTGGCCG 60
61 TTTTGGATGCGGAGATGACCCGCGCCAAAGAGATTCGCGGTGATGCGCAAGAC 120
61 TTTTGGATGCGGAGATGACCCGCGCCAAAGAGATTCGCGGTGATGCGCAAGAC 120
121 AATGTCGATGCGGCGCTGTTACCTTCTTATCACTGATCACTACTATTCCGGCTG 180
121 AATGTCGATGCGGCGCTGTTACCTTCTTATCACTGATCACTACTATTCCGGCTG 180
181 TACTGCTATTTGGAGAGCAAGTATGCGGATGATGATGACCAAGAGAGATTT 240
181 TACTGCTATTTGGAGAGCAAGTATGCGGATGATGATGACCAAGAGAGATTT 240
241 TCGGCGGAGATGACCGCGGCGCCAGCCCTGCGCGGAGCTTGGCGGACATCACTAC 300
241 TCGGCGGAGATGACCGCGGCGCCAGCCCTGCGCGGAGCTTGGCGGACATCACTAC 300
301 ACCGACTGGCGCGCGGACATTTCTATCGCGCGGTGCGGACGCTGACACGCGCGCAAG 360
301 ACCGACTGGCGCGCGGACATTTCTATCGCGCGGTGCGGACGCTGACACGCGCGCAAG 360
361 CGCATGCGGATGAGATTGACCAAGCTGATCTGACCTTCCGCGGAGCTGAGAGAGCC 420
361 CGCATGCGGATGAGATTGACCAAGCTGATCTGACCTTCCGCGGAGCTGAGAGAGCC 420
421 CTACCGGCGGCTGAGATTCTGTCGATCACTACAGCCCTGATGATGATGCGCAATCAAG 480
421 CTACCGGCGGCTGAGATTCTGTCGATCACTACAGCCCTGATGATGATGCGCAATCAAG 480
481 TCGCTGGAAGAGCAAGAGTGTATCGCGAGGCGCGCGGTGTGTGACGTGCGGCGCGG 540
481 TCGCTGGAAGAGCAAGAGTGTATCGCGAGGCGCGCGGTGTGTGACGTGCGGCGCGG 540
541 GCGTCGCGGCGGCTGCGCATCAAGGCGCGCGGTGCGGAGATGAGTGGCCATCCGACAC 600
541 GCGTCGCGGCGGCTGCGCATCAAGGCGCGCGGTGCGGAGATGAGTGGCCATCCGACAC 600
601 AATGCGATGATCCGCGAGATGCGCAAAATGTTCCCTTCTGAGGCTGATGACACTG 660
601 AATGCGATGATCCGCGAGATGCGCAAAATGTTCCCTTCTGAGGCTGATGACACTG 660
661 ACCGCTTCCAGTCCGCGCATCAACACGAGCGGCGGCAATCCGCTCAACACCGCATC 720
661 ACCGCTTCCAGTCCGCGCATCAACACGAGCGGCGGCAATCCGCTCAACACCGCATC 720
721 GTGCAATCGGCGGATCTTTTCTGCTCAACACCTTCCGATGATCTTGGCTACTACAC 780
721 GTGCAATCGGCGGATCTTTTCTGCTCAACACCTTCCGATGATCTTGGCTACTACAC 780
781 GGGCGTGAAGCGGACGCTGTTCTGCGACATGTCATGAGCGGAGCTGACATCTGGGAG 840
781 GGGCGTGAAGCGGACGCTGTTCTGCGACATGTCATGAGCGGAGCTGACATCTGGGAG 840
841 AAGAACTGAGCGGCTGATGCGCGGCGCTGAGCTGATCAAGCGGCGCGCGCTGCAAG 900
841 AAGAACTGAGCGGCTGATGCGCGGCGCTGAGCTGATCAAGCGGCGCGCGCTGCAAG 900
901 GACATCGGCTGATGAGCTCAACAGATGTAACGCGAGTGGGAGCTGTAAGCTGCTCC 960
901 GACATCGGCTGATGAGCTCAACAGATGTAACGCGAGTGGGAGCTGTAAGCTGCTCC 960
961 TTCGCTATGAGCACTCTTCTGCGGCTGCTGTCGACTACTAGGTCGAGGCGCGGCTG 1020

Db 961 TTCGCTATGAGCACTCTTCTGCGGCTGCTGTCGACTACTAGGTCGAGGCGCGGCTG 1020
Qy 1021 GAGCTGCGGAGAGCATGACACACCGAGCTGAGAGCCCGGAGATGATGTTCTTCAAGAGCCG 1080
Db 1021 GAGCTGCGGAGAGCATGACACACCGAGCTGAGAGCCCGGAGATGATGTTCTTCAAGAGCCG 1080
Qy 1081 ATGATGATCTGCTGCGGAGGAGCATGAGCGGCTGCGGCGGCTATGCGAGACGACATCTCTG 1140
Db 1081 ATGATGATCTGCTGCGGAGGAGCATGAGCGGCTGCGGCGGCTATGCGAGACGACATCTCTG 1140
Qy 1141 ATGTCGAGGAGAGACGCTGCCGAGACATCAACCGGCTTCCGTTCCGTCGGAACAAC 1200
Db 1141 ATGTCGAGGAGAGACGCTGCCGAGACATCAACCGGCTTCCGTTCCGTCGGAACAAC 1200
Qy 1201 ATCATCCGCAAC 1212
Db 1201 ATCATCCGCAAC 1212

RESULT 3

US-08-947-726A-1
Sequence 1, Application US/08947726A
Patent No. 5932466
GENERAL INFORMATION:
APPLICANT: Furukawa, Keisuke
APPLICANT: Suzuki, Maroaru
APPLICANT: Koyama, Yasuji
TITLE OF INVENTION: A NOVEL CREATIVE AMIDINOHYDROLASE GENE,
TITLE OF INVENTION: A NOVEL RECOMBINANT DNA, AND A PROCESS FOR PRODUCING
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSES: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/947,726A
FILING DATE: 09-OCT-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/535,444
FILING DATE: 27-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 94/235737
FILING DATE: 29-SEP-1994
REFERENCE/DOCKET NUMBER: 7005-120-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1215 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: AlkaliGenes sp. KS-85
STRAIN: FERM BP-4487

US-08-947-726A-1

Query Match: 99.6%; Score 1207.2; DB 2; Length 1215;
 Best Local Similarity: 99.8%; Pred. No. 8.4e-242;
 Matches 1209; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

1 ATGACTGACGACATGTTGACGATGAAATGACCAACGGGAGAAATTTATTCGCGG 60
 1 ATGACTGACGACATGTTGACGATGAAATGACCAACGGGAGAAATTTATTCGCGG 60
 61 TTTTGGATGCGGAGATGACCGCGCGCAAAAGAGTTGCGGCGTGTGATGGCAAGAC 120
 61 TTTTGGATGCGGAGATGACCGCGCGCAAAAGAGTTGCGGCGTGTGATGGCAAGAC 120
 121 AATGTCGATGCGGCGGCTGTTCACTTCTATCACTGATCACTAATTTCGGCTGCTG 180
 121 AATGTCGATGCGGCGGCTGTTCACTTCTATCACTGATCACTAATTTCGGCTGCTG 180
 181 TACTGCTATTTGCGAGCGAAGTACGCGCATGTCATGACCAACAAGCGCACAGATT 240
 181 TACTGCTATTTGCGAGCGAAGTACGCGCATGTCATGACCAACAAGCGCACAGATT 240
 241 TCGGCGCGGATGACG 300
 241 TCGGCGCGGATGACG 300
 301 ACCGACTGCG 360
 301 ACCGACTGCG 360
 361 CGCATCGGATGAGATTCACCAACGTCATCTGACCTTCGCGCGCGCGCGCGCGCG 420
 361 CGCATCGGATGAGATTCACCAACGTCATCTGACCTTCGCGCGCGCGCGCGCGCG 420
 421 CTACCGGCGGCTGAGTTCGTCGACATCAAGCCCTGATGATGATGCGCACTCAAG 480
 421 CTACCGGCGGCTGAGTTCGTCGACATCAAGCCCTGATGATGATGCGCACTCAAG 480
 481 TCGCTCGAAGAGAGAGAGTTCGTCGACATCAAGCCCTGATGATGATGCGCACTCA 540
 481 TCGCTCGAAGAGAGAGTTCGTCGACATCAAGCCCTGATGATGATGCGCACTCA 540
 541 GCGTGGCGGCTGCGCATCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
 541 GCGTGGCGGCTGCGCATCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
 601 AATGCGATGATCGCGAGATTCGCAATCGCTTCCCTTCTGATGATGATGCGCACTG 660
 601 AATGCGATGATCGCGAGATTCGCAATCGCTTCCCTTCTGATGATGATGCGCACTG 660
 661 ACTGCGTTCAGTTCGCGCATCAACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
 661 ACTGCGTTCAGTTCGCGCATCAACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
 721 GTCGATTCGCGCGCATCAACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 780
 721 GTCGATTCGCGCGCATCAACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 780
 781 GCGCTGAGCGCGCATCAACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 840
 781 GCGCTGAGCGCGCATCAACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 840
 841 AAGAAAGTGGCGCGTTCGCGCATCAACCGAGCGCGCGCGCGCGCGCGCGCGCG 900
 841 AAGAAAGTGGCGCGTTCGCGCATCAACCGAGCGCGCGCGCGCGCGCGCGCGCG 900
 901 GACATGCGCATTCAGTTCGCGCATCAACCGAGCGCGCGCGCGCGCGCGCGCGCG 960
 901 GACATGCGCATTCAGTTCGCGCATCAACCGAGCGCGCGCGCGCGCGCGCGCGCG 960
 961 TTCGCGCATTCAGTTCGCGCATCAACCGAGCGCGCGCGCGCGCGCGCGCGCG 1020
 961 TTCGCGCATTCAGTTCGCGCATCAACCGAGCGCGCGCGCGCGCGCGCGCGCG 1020

1021 GAGCTGCGGAGAGATGACACCGAGCGGAGCGCGCGCGCGCGCGCGCGCGCG 1080
 1021 GAGCTGCGGAGAGATGACACCGAGCGGAGCGCGCGCGCGCGCGCGCGCGCG 1080
 1081 ATGCTGATGCTGCGGAGAGATGACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1140
 1081 ATGCTGATGCTGCGGAGAGATGACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1140
 1141 ATGCTGAGGAGAGAGATGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1200
 1141 ATGCTGAGGAGAGAGATGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1200
 1201 ATCATTCGCGCAC 1212
 1201 ATCATTCGCGCAC 1212

RESULT 4

US-09-252-991A-4171
 ; Sequence 4171, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; PRIOR FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 4171
 ; LENGTH: 858
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-252-991A-4171

Query Match: 6.2%; Score 74.8; DB 4; Length 858;
 Best Local Similarity: 45.7%; Pred. No. 1.4e-06;
 Matches 339; Conservative 0; Mismatches 397; Indels 6; Gaps 2;

471 CACCATCAAGTCTGCGAAGAGAGAGATGATCGCGAAGGCGCGCGCGCGCGT 530
 81 CACCATCAAGAGAGAGAGATGATCGCGAAGGCGCGCGCGCGCGCGCGCGCG 140
 531 CGGCG 590
 141 AGTGTGAAATGATCGCGAAGAGAGATGATCGCGAAGGCGCGCGCGCGCGCG 200
 591 CGGCG 650
 201 CATGCG 260
 651 GGAACCTGAGATCTGTTCCATCGCGGATCAACCGAGCGCGCGCGCGCGCGCG 710
 261 GGGCTTTCGAAGTGTGATCGATCGATCAACCGAGCGCGCGCGCGCGCGCGCG 770
 711 CAACGCGATCGTCAATCGCGGAGATCTGTTGCTCAACCGATCGCGCGCGCG 770
 321 CGAAGAGCG 380
 771 CTACTACACCG 830
 381 CTACACCG 440
 831 CATTCG 890
 441 CCTTCG 500
 891 GCGTGCAGAGAGATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 950

```

Db      501 GCACCTGCGCGGATATCG---GCGAAATCATTCAGAAAGCAGCGAAAGAAAGCGCTTCTC 557
Qy      951 GTACCGCTCTTGGGTATATGACCATCTCTTGGCGGTCTGTGCATCTACTACAGGTGCCA 1010
Db      558 GGTGTCCCGGAATCTATGCGCGCCATGCGCATTCGGCAAGGTCTTC---ACGAGAGCGCGA 614
Qy      1011 GCGCGCGGTGAGCTCGCGGAGACATCGACCGAGCTGAAGCCGGCATGTGTCTTC 1070
Db      615 GGTCTTCACACTACGGTCTGTGCGGGTACCGGCATCGAGCTGAAGAAAGCATATCTTCAC 674
Qy      1071 CATGAGCCGATGTGTATGTCTGCGGAGGCGCATGCCGTGTGCGGCTATTCGAGCA 1130
Db      675 CATCGAGCCGATGTATCAACAGAGGCGCGCGGAGACCGCGCTGTCTGCGGACGGCTGAC 734
Qy      1131 CGACATCTCTGATGTGTGGGAGAGACGGTCCCGAATCATCCGGCTTCCCTTGGGTGC 1190
Db      735 CGCGATCACCAAGAGCCGCAAGCTGTCCGCGCATGTGAGAGCACACGTGTGTCTCACCG 794
Qy      1191 GGAACACAAATCATCTCCGCAAC 1212
Db      795 CGATGGCTTACGAGATCTCTCAC 816

```

RESULT 5

```

US-09-252-991A-4435/c
; Sequence 4435, Application US/09252991A
; Patent No. 6551795

```

GENERAL INFORMATION:

```

; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 4435
; LENGTH: 963
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-4435

```

```

Query Match      6.2%; Score 74.8; DB 4; Length 963;
Best Local Similarity 45.7%; Pred. No. 1.4e-06;
Matches 339; Conservative 0; Mismatches 397; Indels 6; Gaps 2;

```

```

Qy      471 CACCATCAAGTCTCGAAGAGACAGATGATCCGCGAAGCGCCGCTGTGTGAGCT 530
Db      922 CACCATCAAGAGCGCCGAGATATCGAAGATGCGCATGCGCGCGCTGCGCGCGCA 863
Qy      531 CGCGCGCGCGCTGCGCGCTGCGCATCAAGCGCGCGCGCGCGCATGAATGGCGAT 590
Db      862 AGTCTGGAATATGATGCGGAACAATCAAGCCCGCGGAGACCAAGAGACTGATG 803
Qy      591 CGCCACCAACATGAGATATCGCGAGATCCGCAATATGTTCCCTTGTGAGCTGAT 650
Db      802 CATCTGCCACGACTATATGTGTCAAGAGAGAGAGGATTCGCCGCCCTTGAATCA 743
Qy      651 GGAACCTGAGACTGTCTCACTGCGGAGATCAACCGAGCGCGCGCATATCCGCTAC 710
Db      742 GGGCTTTCGAAGTGTGATGCACTCGATCAACATGTGTCTGCGATGGCATCCCA 683
Qy      711 CAACCGCATCTGAGCATCGCGGAGATCTTTCGCTCAACATCTTCGCGATGATCTTGG 770
Db      682 CGAGAGCGCGCTGAAGAGAGGAGACATCTCAACGTCGACATCACTGATCAAGAGCG 623
Qy      771 CTACTACACCGCGCTGAGCGGACCGCTGTTCTGCGACATGTGATGACCGCAGCTCGA 830
Db      622 CTACAGCGCGGAGACAGAGAAATGTTCTGCTGTGCGCAAGACCGGAAATGGCGCGACCG 563

```

```

Qy      831 CATCTGAGAGAGAAAGTGTGCGCGCATGCGCGCGGCTCGAGCTATCAAGCGCGGCGC 890
Db      562 CTTCTGCCAGATCAACCGAGATGATGATCAAGGCAATTTCCGTGTGCTGCTCGGCGC 503
Qy      891 GCGCTGCAAGACATGCCATGAGCTCAACAGATGTAACCGAGTGGAGCTGTGAA 950
Db      502 GCACTGGGCGATATCG---GCGAATCATTCAGAAAGCAGCGGAAAGAAAGCGCTTCTC 446
Qy      951 GTACCGCTCTTGGGTATATGACCATCTCTTGGCGGTCTGTGTGCACTACTACGTGCCA 1010
Db      445 GGTGTCCCGGAATCTGTGGCGCATGCGCATGCGCAAGGTCTTC---ACGAGAGCGCA 389
Qy      1011 GCGCGCGGTGAGCTGCGCGAGACATCGACACCGAGCTGAAGCCGGCATGTGTCTC 1070
Db      388 GGTCTTCACACTACGGTCTGTGCGGGTACCGGCATCGAGCTGAAGAAAGCATATCTTCAC 329
Qy      1071 CATGAGCCGATGTGTATGTCTGCGAGGCGCATGCCGTGTGCGGCGCTATTCGAGCA 1130
Db      328 CATCGAGCCGATGTATCAACAGAGCGCGCGCGAGACCGCGCTGTCTGCGGACGGCTGAC 269
Qy      1131 CGACATCTCTGATGTGTGGGAGAGACGGTCCGAGAAATCATCCGGCTTCCCTTGGGTGC 1190
Db      268 CGCGATCACCAAGAGCCCGAAGCTGTCCGCGCATGTGAGAGCACACGTGTGTCTCACCG 209
Qy      1191 GGAACACAAATCATCTCCGCAAC 1212
Db      208 CGATGGCTTACGAGATCTCTCAC 187

```

RESULT 6

```

US-09-252-991A-4314
; Sequence 4314, Application US/09252991A
; Patent No. 6551795

```

GENERAL INFORMATION:

```

; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 4314
; LENGTH: 1176
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-4314

```

```

Query Match      6.2%; Score 74.8; DB 4; Length 1176;
Best Local Similarity 45.7%; Pred. No. 1.5e-06;
Matches 339; Conservative 0; Mismatches 397; Indels 6; Gaps 2;

```

```

Qy      471 CACCATCAAGTCTCGAAGAGCAAGATGATCCGCGAAGCGCCGCTGTGTGAGCT 530
Db      244 CACCATCAAGAGCGCCGAGATATCGAAGATGCGCATGCGCGCGCTGCGCGCGCA 303
Qy      531 CGCGCGCGCGCTGCGCGCTGCGCATCAAGCGCGCGCGCGCGCATGAATGGCGAT 590
Db      304 AGTCTGGAATATGATGCGGAGAACATCAAGCCCGGTGACACCGAGACTGATG 363
Qy      591 CGCCACCAACATGAGATATCGCGAGATCCGCAATATGTTCCCTTGTGAGCTGAT 650
Db      364 CATCTGCCACGACTATATGTGTCAAGAGAGAGGATTCGCCGCCCTTGAATCA 423
Qy      651 GGAACCTGAGACTGTGTTCAGTGTGCGGATCAACCGAGCGCGCGCAATCCGCTAC 710
Db      424 GGGCTTTCGAAGTGTGATCTGACCTCGATCAACCATGTGTCTGCGCATGTGCCCA 483
Qy      711 CAACCGCATCTGAGCATCGCGGAGATCTTTCGCTCAACATCTTCGCGATGATCTTGG 770

```


Db 1462 CTCGACGAGGACCTCTGGGCTTCGTGACCTGATCCCGAGCGCGGCGAGCGGAT 1521
Qy 897 CAGAGGACATGCGCATGAGCTCAACGAGATGTAACCGAGTGGAACTCTGAAGTACCG 956
Db 1522 CATCGACATCGCTCGACCGAGATTCGCGGACGCGCTCGGCTACCAACAGTACCGCGCT 1581
Qy 957 CTCCTTGGCTATGCGACCTCTTGGCGGTGTGTGCACTACTAGTGTGCGAGGCGG 1016
Db 1582 CACGAGCGCGGAGACAGACATCGGCTCGGCTTCAACGACGACCGCTGTGGCTCAT 1641
Qy 1017 CGTGAAGCTGCGGAGAGACATCGACACGAGCTGAAGCCCGGATGTGTCTCATGGA 1076
Db 1642 CGGCGGCGTGGCGGCTCATCAAGAGTCCGCGCATCGGCGCATCTCGAGAGCCGCT 1701
Qy 1077 GCGGATGTGATGCTGCGGAGGCGCATGCCGCTGCGCGGCTATTCGAGAGACATCAT 1136
Db 1702 GCGGTTGCAACAGAGCCCGGCTCGAGGTCCTGTTGAGCACTGACCGGCTCTT 1761
Qy 1137 CGTGAATGTCGCGGAGAGCGGTGCGAGAACATCACCGGCTTCCGTTGCTCGGAGAA 1196
Db 1762 CAGTTTACGCTGACAGACCGCGCCGCGACGCGCTGCGCTCATGCGCTGCGAGCTG 1821
Qy 1197 CAACATCATCCGCAAC 1212
Db 1822 GAAAGCTGCTCAAC 1837

RESULT 8
US-09-025-691-2
; Sequence 2, Application US/09025691
; Patent No. 6069299

GENERAL INFORMATION:

APPLICANT: Broadway, Roxanne M.
APPLICANT: Hartman, Gary E.
TITLE OF INVENTION: FUNGUS AND INSECT CONTROL WITH
TITLE OF INVENTION: CHITINOLYTIC ENZYMES
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
STREET: Clinton Square, P.O. Box 1051
CITY: Rochester
STATE: New York
COUNTRY: U.S.A.

ZIP: 14603

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/025,691

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Goldman, Michael L.
REGISTRATION NUMBER: 30,727
REFERENCE/DOCKET NUMBER: 19603/20120

TELECOMMUNICATION INFORMATION:

TELEPHONE: (716) 263-1304
TELEFAX: (716) 263-1600

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1294 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: CDNA
US-09-025-691-2

Query Match 5.3%; Score 64; DB 3; Length 1294;
Best Local Similarity 42.6%; Pred. No. 0.00027;
Matches 390; Conservative 0; Mismatches 522; Indels 3; Gaps 1;

Qy 243 GCGCGGACATCGACGCGGCGCAGCCCTGGGCGCGGAGCTTGGGCAACAATCACTTACAC 302
Db 197 GGTATGACACTTACCCCTTTCGACACCGGCGCGGCGCCCTTCAAGCCCTTCAACCG 256
Qy 303 CGACTGCGCGCGGCAAAATTTCTATCGCGCGCTGCGCGAGCTGACCAACGCGGCGCAAGCG 362
Db 257 CAGGGGCGCGGCTGACCGCGCGCTTCAACCGCGCGCGCTTCCGCGCGCTTCCGCGCTGCGCT 316
Qy 363 CATCGGCATCGAGTTTCAGACACGTCATCTTCCGCGCGGCTTCCGCGCGCTTCCGCGCGCT 422
Db 317 CACCGCGCGCGGAGCGCGCGCGGCGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 376
Qy 423 ACCGCGCGCTGAGTTTGTGACATGACGACGCGCTTCAATGTGATGCGGACCATCAATGTC 482
Db 377 ACCGCGCGCGGAGCGGCTGACCGGTTTACGAGAACTTCAACAACGCGCGCGCGCTGACAGC 436
Qy 483 GCTCGAAGAGCAAGATGATTCGCGAAGCGCGCGCTGTGTGACGTTGCGCGCGCGCGCG 542
Db 437 CCGCGCGCGCGGAGCGCGCGCTTACGATCATGCGCGCTTCTTCCGCGCGCGCGCGCGCG 496
Qy 543 CTGCGCGCGCTGACATCAAGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 602
Db 497 CAACGCGCGCGAGATCACTTCAACCTCGACTCGGTGCGCGCTGCGCGCGCGCGCGCGCG 556
Qy 603 TCGGATGATTCGCGCGAGATCGCAAAATCGTTCCCTTTCGCGAGCTGATGGAACCTGAGC 662
Db 557 GCGATTCGCGCGCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 616
Qy 663 CTGATTTCAGTGGGACATCAACCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 722
Db 617 CGGCGCGCGGAGAGCG 673
Qy 723 GCAATCGCGCGACATCTTGTGCTCAACCTTCCGATGATCTTCCGCTACTTCAACCGC 782
Db 674 CGACGACACTTACGCGCTGATGAGAGTACGCGCTTTCGACGCGCTGACATGACCTGGA 733
Qy 783 GCTGAGGCGGACGCGCTTCTGCGACCATGTGATGAGCGCGCGCGCGCGCGCGCGCGCG 842
Db 734 GAAAGGCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTT 793
Qy 843 GAAAGGCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTT 902
Db 794 GAAAGGCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTT 853
Qy 903 CATGCGCATCGAGCTCAACGAGATGTACCGGAGTGGACCTGCTGAAGTACCGCTCTT 962
Db 854 GTACTTCAAGACGCGCGCTGTCAAGAAAGACTTCTGACCGCGCGCTCAACATGACGACTGA 913
Qy 963 CGGCTATGCGCATCTTCTGCGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1022
Db 914 CAACGCGCGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 973
Qy 1023 GCTGCGGAGAGACATCGACACGAGCTGAACCGCGCGCATGTGTCTTCATGAGACCGAT 1082
Db 974 CTTCAACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1033
Qy 1083 GGTATGCTGCGGAGGAGCATGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1142
Db 1034 CGGTGCTGCG 1093
Qy 1143 CGTGGGAGAGAGCG 1157
Db 1094 CGAGCGCTGAGCTG 1108

RESULT 9

US-09-902-540-1206/C
; Sequence 1206, Application US/0902540
; Patent No. 6833447

GENERAL INFORMATION:

APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.

APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15849)B
CURRENT APPLICATION NUMBER: US/09/902,540
PRIOR APPLICATION NUMBER: 2001-07-10
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
LENGTH: 18686
TYPE: DNA
ORGANISM: Myxococcus xanthus
US-09-902-540-1206

Query Match 5.3%; Score 63.8; DB 4; Length 18686;
Best Local Similarity 44.1%; Pred. No. 0.00046;
Matches 407; Conservative 0; Mismatches 507; Indels 9; Gaps 3;

242 CGGCGGCGATCGACGGCGCCAGCCCTGGCGCGGAGCTTGGCGGACCAATCACCCTTACA 301
13107 CGCCCGCGGTGCTTCCGGGCCCCCAAGACACCTGCTGTGTGGCAGTGGCACAAGTACA 13048
302 CCGACTGGGCGCGCCGACATTTCTATCGGCGCGTGGCCAGCTGACCAAGCGCGGCAAGC 361
13047 AGAAGTGCACCGCGCGCGGATGCGCGCGCGGCAAGAGCGCCGCGAGGTCTTC 12988
362 GCATCGGCGATCGAGTTGACACGTCATCTGATCTTCCGCGCGACGTCGAGAGCC 421
12987 GCAAGGCGATCGCGCGCGCGCTGATACGCCCGCGCGAGGTGCGCTCCACATCCCC 12928
422 TAACCGGCGCTGAGTTGTCGACATCAGCAGCCCTCGATGTGATGCGCACCATCAGT 481
12927 GTCCGAGCTACGCGCTGACGCGCGCGCGCTCGCGCAAGAGCCGCGCTCCGACATCAGA 12868
482 CGCTGAGAGAGAGAGAGTGAATCCGCGAGAGCGCGCGCGCGCTGTGTGACGTGCGCGCG 541
12867 CGCCGAGCGCTCATCG 12808
542 CCGCGCGCGCTGATCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 601
12807 AGGTGCGCTCCACG 12748
602 ATGCGATGATCGCGCGAGATCGCGCAATCGTCCCTTGTGTGAGCTGATGACACCTGGA 661
12747 AGCGCGATCAAGCG 12688
662 CCGTGTTCAGTTCGGGATCAACACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 721
12687 CGCTCTGACGTCGCGTCAAGAGGTATCTGCGACGCGCATCCCGA--CAGCGCGCGCG 12631
722 TGCATTCGCGCGCATCTTGGCTCAACACCTTCCGATGATCTTGGCTACTACACCG 781
12630 TGAAGAGCGCGATATCTGCAACTGAGTCAACATCTGATGATGATGCGAGCG 12571
782 CGCTGAGCGCACGCTGTTCTGCGACATGTCGATGACCGCAGCGCTCGACATCTGGAGA 841
12570 ACTGCTCGCGAGTACTTCTGTGGGCAAGTGAAGAGAGCAGCGCGTGTGCGCG 12511
842 AGAAGTGGCGCGTCACTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 901
12510 TGAAGCGCGAGTGCCTGAGCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 12451
902 ACATCGCGATCGAGCTCAAGAGATGTAACGCGAGTGGAGACCTGCGAGTACCGCTCT 961
12450 ACATCGCGCGCGCGCGATGAGAGCGATGCAACCGAGCGCGATGACCGTGTGCGCGCT 12391
962 TCGGCTATGAGCACTCTTTCGCGGTGCTGTGCACTAATCGTTCGCGAGCGCGCGGTG 1021
12390 ACTGC--GCGCAGCGCGATCGGAGAGCGTTCCACACTTCGCTCAGGTGCGCGACTACT 12334
1022 AGCTGCGCAGAGCATCGACCGGAGCTGAAGCGCGCGCGCGCGCGCGCGCGCGCG 1081

DB 12333 ACGAGCGCGAGG---CCGACACCGTCAATGACGCCCGCGCGATGTTTCAACCGTGCAGCCGA 12277
QY 1082 TGTGATCTCCGAGGCGCATGCCCGGTGCGCGCGCTATCGGAGCAGACATCTTGA 1141
DB 12276 TGATCAACTGAGGCGCGCGGTGATGAGAGTGAAGTGAACCGCGCGTACCG 12217
QY 1142 TCGTGGGAGAGAGCGTGCAGG 1164
DB 12216 CCGACGCGACCGCGCAGCGCCGAG 12194

RESULT 10
US-08-343-428-1
Sequence 1, Application US/08343428
Patent No. 565586
GENERAL INFORMATION:
APPLICANT: Nakamura, Etsuo
APPLICANT: Tezuka, Hiroshi
APPLICANT: Kitadokoro, Kengo
APPLICANT: Shin, Naoharu
APPLICANT: Teraoka, Hiroshi
TITLE OF INVENTION: No. 565586e1 Protease
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESS: James F. Haley, Jr., Fish & Neave
STREET: 1251 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10020-1104
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 720KB storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS Dos 5.0
SOFTWARE: WordPerfect 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/343,428
FILING DATE: 18-NOV-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/J93/00592
FILING DATE: 30-APR-1993
APPLICATION NUMBER: JAPAN 4-126511
FILING DATE: 19-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Haley, Jr., James F.
REGISTRATION NUMBER: 27794
REFERENCE/DOCKET NUMBER: SHN-7
TELEPHONE: (212)596-9000
TELEFAX: (212)596-9090
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2064
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
ORIGINAL SOURCE:
ORGANISM: Streptomyces fradiae
FEATURE:
NAME/KEY: -35 signal
LOCATION: 359..364
IDENTIFICATION METHOD: by experiment
NAME/KEY: -10 signal
LOCATION: 378..383
IDENTIFICATION METHOD: by experiment
NAME/KEY: CDS
LOCATION: 435..1505
IDENTIFICATION METHOD: by experiment
NAME/KEY: sig peptide
LOCATION: 435..944
IDENTIFICATION METHOD: by experiment

US-08-343-428-1

Query Match 5.1%; Score 62; DB 1; Length 2064;
 Best Local Similarity 44.4%; Pred. No. 0.00075;
 Matches 342; Conservative 0; Mismatches 420; Indels 8; Gaps 2;

```

QY 103 GCGTGGATGCGCAAGAACATATGTCATGCGGCGCTTTGCACTTATCATGATCAAC 162
DB 993 GCTTCAAGTCAACCAAGAACGCGCTTCTGTAATCTTCAACCGCGGCACTGCAAC 1052
QY 163 TACTATTCGCGCTGCTGTAATGCTATTTTGGACGCAAGTACGGAGTGTATGACAC 222
DB 1053 CTCTGTCACCTGCTGCTGCTCAAC---TCGGGGGACGCTTCATCGGCTCGCGAGGAC 1109
QY 223 AACAAAGCAGCAATTTTCGCGCGGCTGCAAGGCGGCAAGCCCTGGCGCGGCACTTC 282
DB 1110 ACCACCTTCCGACCAACGACTACGCGCATGCTCCGTTACACCAACGACGACGTAAC 1169
QY 283 GCGCAACATCATCTACACGACTGGCGCGGCAATTTTATGCGCGGCGGCGACG 342
DB 1170 GCGCGGCTCACTGTATCAACGCGCGCTACAGACATGCTCCGCGGCGGCGGCGCTC 1229
QY 343 CTGACCAAGCGGCGCAAGGCAATCGGCAATGAGTTGACACGTCATCTGCACTTCCG 402
DB 1230 GTGGGCGAGGCGCATCAAGAAAGCGGCTCCAGACCAAGTCAACGAGGCGGCACTG 1289
QY 403 GCGCAGCTGAGAAAGCCTTACCGGCGCTGCAAGTTGCTGACATGACGAGCCCTGATG 462
DB 1290 GCGCTCAACGTCACGCTCACTACAGCGACGCGCGGCTTACGGAATGCTGCGACAG 1349
QY 463 TGGATGCGGCAATCAAGTGCCTGCAAGAGCAAGAGCTGACGCGAAAGCGCGCGCTG 522
DB 1350 GCGTCTCGCGCGGCGGCGCAAGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1409
QY 523 TGTGACGTGCGGCGGCGGCGCTGCGCGCTGCTCAACGAGCGGCGGCGGCGGCGGCG 582
DB 1410 ATCACTGCGGCGAGCTCGGCTGCAAGGCGCAACAGGCTCCGCTACCAAGCGGCTG 1469
QY 583 GTGGGATGCGCAACCAATGATGATGCGGAGATGCGCAATGCTTCCCTTCTG 642
DB 1470 GCGGAGGCGCTGAGCGCTATGCGCTGAC-----GTGATGACGCGCGCGCGCGG 1524
QY 643 GAGCTGATGACACCTGGAAGCTGTTTCAAGTGGGATCAACAGCGGCGGCGGCAAT 702
DB 1525 GCGCGAGTGCACAGTGGCGGATGATGATGATGCTCTCATACCGGCGGCGGCGGCG 1584
QY 703 CCGGTCAACCAACCGGATGCTGCAATCGGCGCAATCTTTCGCTCAACCTTCCGATG 762
DB 1585 CTGGCAAGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1644
QY 763 ATCTTGGCTACTACACGCGCTGAGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTG 822
DB 1645 CTGAGCGGAGGAGAGTGGGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1704
QY 823 AGCCTGCAATCTGCGGAGAGAAAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 872
DB 1705 GCGCGGCGGAGTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1754

```

RESULT 11
 US-09-475-515-54
 Sequence 54, Application US/09475515A

GENERAL INFORMATION:
 Patent No. 6602705
 APPLICANT: BARNETT, Susan
 APPLICANT: ZUR MEGEDE, Jan
 APPLICANT: SRIVASTAVA, Indresh
 APPLICANT: LIAN, Ying
 APPLICANT: HARTOG, Karin
 APPLICANT: LIU, Hong
 APPLICANT: GREER, Catherine
 APPLICANT: SELBY, Mark
 APPLICANT: WALKER, Christopher

TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
 OF VIRUS-LIKE PARTICLES
 FILE REFERENCE: 1621.002
 CURRENT APPLICATION NUMBER: US/09/475,515A
 NUMBER OF SEQ ID NOS: 90
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 54
 LENGTH: 1599
 TYPE: DNA
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Description of Artificial Sequence: gp120.modus4
 US-09-475-515-54

Query Match 5.0%; Score 60.6; DB 4; Length 1599;
 Best Local Similarity 42.6%; Pred. No. 0.0014;
 Matches 434; Conservative 0; Mismatches 579; Indels 6; Gaps 2;

```

QY 196 CCGAATAGCGATGTCATGCAACCAACGCGGCAAGCATTTTCGCGCGGCGGCGGCGG 255
DB 364 CCGTGTGCTGACCTGTAACGCAACGAGCTGACCGGCGGCAACCGGCGGCGGCGG 423
QY 256 GCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 315
DB 424 AGCACCAAGCGGCAACCAAGCAACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 483
QY 316 GACAAATTTTATGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 372
DB 484 AGCTGGAGAAAGATCCCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 543
QY 373 GAGTTGACCAAGTGAATCTGACTTTCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 432
DB 544 GTGGCGCAAGAGTGCAGAGAGTACAGCTGTTTCAAGCTGAGAGTGTGCTGCTATC 603
QY 433 GAGTTGTCGACATCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 492
DB 604 GACAAAGCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 663
QY 493 CAGAAAGTATCCGGAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 552
DB 664 TGCCCAAGGTGAGGTTGAGAGCCCATCCCATCTCTGCGGCGGCGGCGGCGGCGG 723
QY 553 GCATCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 612
DB 724 ATCTGGAAGTCAAGGAGCAAGAGTCAAGGCGGCGGCGGCGGCGGCGGCGGCGG 783
QY 613 GCGGAGATCGGCAATGCTTCCCTTGTGAGGCTGATGACACCTGGAAGCTGTTGAG 672
DB 784 GTGCAATGACCAAGGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 843
QY 673 TCGGCGATCAACCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 732
DB 844 CTGGCGGAGGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 903
QY 733 GACATCTTTTGTCTCAACCTTCCCATGATCTTCCGCTACTACACCGCGTGAAGCG 792
DB 904 ATCTGCAAGTGAACGATGCTGAGAGTCACTGCAATCGGCGGCGGCGGCGGCGG 963
QY 793 ACGGTTTCTGCGACATGTCATGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 852
DB 964 AAGAGATCAACATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1023
QY 853 GTGATGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 912
DB 1024 ATCGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1083
QY 913 GAGCTCAACGAGATGACCGGAGTGGGAGCTGCTGAGATACCGTCTTGGGCTATGCG 972
DB 1084 G---TGAGAGAGTGGCGGAGGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1140
QY 973 CACTCTTTCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1032

```

```

Db      1141 AGCGCGGCGAGCCCGAGATCGTGTTCACAGCTTCACTGCGCGCGAGATTCTTCTAC 1200
Qy      1033 GACATCGACACGAGCTGAAGCCCGGATGAGTGTCTCCATGAGAGCCGATGATGCTG 1092
Db      1201 TCCACACACGAGCAGCTGTTCACAGCAGCTGGAAATCATCAGGAGGATGAACAAGACC 1260
Qy      1093 CCGAGAGGAGATCCCGGCTGCTGCGGAGATTCGCGAGACGACATCTGTATCTCGGAGAG 1152
Db      1261 AAGGAGAAAGACACATCATCTCTGCGCTCCGATCCGCAATCCGATCAATCATGTGGCAG 1320
Qy      1153 GACGATGCGGAGACATCAACCGGCTTCCGTTCCGTCGCGGAAACAACATCATCCGCA 1211
Db      1321 GAGGTGGGAGGAGGATGATGACGCCCCCCCCCATTCGCGGCGACATCAAGTGCAGACGAA 1379

```

RESULT 12
US-09-475-515-56
Sequence 56, Application US/09475515A

```

; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 56
; LENGTH: 2112
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: gp140.modus4
US-09-475-515-56

```

Query Match 5.0%; Score 60.6; DB 4; Length 2112;
Best Local Similarity 42.6%; Pred. No. 0.0015;
Matches 434; Conservative 0; Mismatches 579; Indels 6; Gaps 2;

```

Qy      196 CCGAAGTACGGATGTATTCACCAACAACGCAACGATTTCCGCGCGCATCCAG 255
Db      364 CCCCTGTGTGATACCTGTGATGCAACGCAAGCTGACCGGCGACCAAGCGCACAAAC 423
Qy      256 GCGGCGCAGCCCTGGCGCGCAGCTTCCGCGACCAATCATCACTACAGCATGGCGCGC 315
Db      424 AGCACCAGGAGGACCAACAGCAGCAGCGGACCAACAGACCAAGCAGCAGCAGC 483
Qy      316 GACCAATTTCTATCGCGCGCTGCGCGACGT--GACCAAGCGCGCGCAAGCGCATCGGCATC 372
Db      484 AGCTGGAGAAAGATGCCAGGCGGAGATCAAGAACTGCAAGCTTCAATCACTACACAGC 543
Qy      373 GAGTTGACCAAGTCAATTCATCTCCGCGCGCAGCTGAGAGAAAGCCCTTACCGGCGTC 432
Db      544 GTGCGCGAAGAGTGAAGAGATGACAGCTGTCTTAACAGCTGACGTGTGCCATC 603
Qy      433 GAGTTGTCGATCATCAACGCGCTTCATGTGATGATCGCAACATCAATCGCTCGAAGAG 492
Db      604 GACCAAGCAAGACGCGCTACCGCTGATCACTGCAACACAGCTGATCAACCGAGCC 663
Qy      493 CAGAAAGTATCCGCGAGCGCGCGCTGTGTGACGTGCGCGCGCGCGCTGCGCGCT 552
Db      664 TGCACCAAGGTAGCTTCAAGCCCATCCCATCACTACTGCGCGCGCGCGCTTCCGC 723
Qy      553 GGCATCAAGGCGCGCGCTGCCAGATGAAGTGGCATGCGCACCAACCAATCGATGATC 612

```

```

Db      724 ATCTGAAGTGCAGAGACAAAGATTTCACGAGCACCGGCCCTCTGCAAGACGTGAGCAC 783
Qy      613 CGCGAGATGCGCAAAATCTTCCCTTCTGTGAGCTGATGACACCTGACCTGTTCCAG 672
Db      784 GTGAGTGCACCAAGGATCCGCGCGTGTGAGCACCCAGCTCTGCTGTAACGGCAGC 843
Qy      673 TCGGCGATCAACACCGAGCGCGCGCAATCCGCTACCAACCGCATGTGTGAATCCGCG 732
Db      844 CTGCGCGAGAGAGATGTCTGTGCTCCAGAACTTACCGAACAACGCCAAGACATC 903
Qy      723 GACATCTTGTGCTCAACACCTTCCGATGATCTTCCGCTACTACACCGCGCTGAGCGC 792
Db      904 ATCTGACGCTGACAGATCTGCTGAGATTCATCTGATCTCCGCCCAACAACAGCGCT 963
Qy      793 AGCTGTTCTGCGACATGTGATGACGCGCAGCTTCGACATCTGAGAGAAAGACGTGCC 852
Db      964 AAGAGCATTCACATGCGCGCGCGCGCTTCTACGCGCACCGGCGACATCATCGGCGAC 1023
Qy      853 GTGATTCGCGCGCGCTGAGCTGATCAAGCCCGCGCGCGCTGCAAGGACATCGCCATC 912
Db      1024 ATCCGCGAGGCCCATCTGCAACATGCAAGGCAATGCAACCAACCTTCAAGCATC 1083
Qy      913 GAGCTCAAGAGATGTACCGGAGTGGAGCTGTGAAGTACCGCTCTCCGCTATGAGC 972
Db      1084 G---TGAGAACTGCGGAGACGTTGGCAACAAAGACATCATCTTCAACAGCAGC 1140
Qy      973 CACTCTTTCGCGCTGCTGTCACCTACTACGATGCGCGAGCGCGCTGAGAGCTGCGCGAG 1032
Db      1141 AGCGCGCGGAGCCCGAGATGATGATGTTCCACAGCTTCACTGCGCGCGGAGATTTCTTAC 1200
Qy      1033 GACATCGACACGAGCTGAAGCCCGGATGATGATGATGATGATGATGATGATGATGATG 1092
Db      1201 TCGAACACACGACGCTGTTCACAGCAGCTGGAACATCACGAGAGGTGAACAAGACC 1260
Qy      1093 CCGAGAGGATGCGCGGCTGCGCGCGGCTATGCGGACGACGATCTGATCTGCGGAG 1152
Db      1261 AAGGAGAAAGACACCATCATCTCTGCTGCGGAGATCCGCGAGATCAAGATGGGAG 1320
Qy      1153 GACGATGCGGAGAAATCATCAACCGGCTTCCGTTCCGTCGCGGAAACAACATCATCCGCA 1211
Db      1321 GAGGTGGGAGGAGGATGATGACGCCCCCCCCCATTCGCGGCGACATCAAGTGCAGACGAA 1379

```

RESULT 13
US-09-475-515-57
Sequence 57, Application US/09475515A
Patent No. 6602705
GENERAL INFORMATION:
APPLICANT: BARNETT, Susan
APPLICANT: ZUR MEGEDE, Jan
APPLICANT: SRIVASTAVA, Indresh
APPLICANT: LIAN, Ying
APPLICANT: HARTOG, Karin
APPLICANT: LIU, Hong
APPLICANT: GREER, Catherine
APPLICANT: SELBY, Mark
APPLICANT: WALKER, Christopher
TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
FILE REFERENCE: 1621.002
CURRENT APPLICATION NUMBER: US/09/475,515A
CURRENT FILING DATE: 1999-12-30
NUMBER OF SEQ ID NOS: 90
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 57
LENGTH: 2112
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:
OTHER INFORMATION: gp140.mut.modus4
US-09-475-515-57

Query Match 5.0%; Score 60.6; DB 4; Length 2112;
 Best Local Similarity 42.6%; Pred. No. 0.0015;
 Matches 434; Conservative 0; Mismatches 579; Indels 6; Gaps 2;

QY 196 CGCAAGTACGGCATGTGATGACCAACCAACGCCAGATTTGGCCGGCATGAC 255
 DB 364 CCCCTGTGCTGACCCCTGAACTGACCGACAGTACCGGACCAACCGGACCAAC 423
 QY 256 GGGGGCCAGCCCTGGGCGCCGAGTTTGGGCAACAATCACTACACCGTGGCCGC 315
 DB 424 AGACACCGCGCAACCAACGACCAAGGCGGACCAACGACCAACGACGACGAC 483
 QY 316 GACAAATTTATGCGCGCGCGCGCGAGCT---GACGACGGGCGCGCAAGCGCATCGGATC 372
 DB 484 AGCTGGAGAGAGATGCCCGAGGGGAGATCAAGAACTGCAAGCTTCAATCAACCAACGAC 543
 QY 373 GAGTTGACCAACGTCATCTGCACTTCCGCGCGAGCTCGAGGAAAGCCCTACCGGCGCTC 432
 DB 544 GTGCGGACAAAGGTGAGAGAGATACAGCTGTTTCAAGAGTGAAGCTGAGCTGAGCCATC 603
 QY 433 GAGTTGTGACATCAAGCCGCTGATGTGATGCGGACCAATCAAGTCTCGAAGAG 492
 DB 604 GACAAAGCAACCGCAGCTACCGCTGATCACTGCAACACAGCGTATCAACCCAGGCC 663
 QY 493 CAGAACTGATCGCGAAGGCGCGCGCTGTGTGAGTGGCGGCGCGCGCTGCGGCGCT 552
 DB 664 TGCCCCAAGGTGAGCTTCAAGCCATCCCATCACTACCTGCGCCCGCGCTTGGCC 723
 QY 553 GCCATCAAGGCGCGCTGCGCGAGATGAAAGTGGCATGCGCACCAACCAATGCGATGATC 612
 DB 724 ATCTGAAAGTGAAGGCAAGAAAGTTCAACGCAACGCGCCCTGCAAGAAAGTGAAGAC 783
 QY 613 CGGGAATGCGCAATCTTCCCTTGTGTGAGTGAAGTGAACCTGGAACCTGTTCCAG 672
 DB 784 GTGCAATGCAACCCAGCGCATCCCGCGGTGTGAGGACCCAGCTGTGTGAACGCGCAG 843
 QY 673 TCGGGCATCAACCGAGCGCGGCAATTCGGTCAACCAAGCGCATGCGCATCGCG 732
 DB 844 CTGGCGAGAGAGATCTGTGCTGCGCTCGAGAACTTCAACGCAACCGCAAGACATC 903
 QY 733 GACATCTTCTGCTCAACACTTCCGAGATCTTGGGCTACTACACCGCGCTGAGAGCG 792
 DB 904 ATCTGACACTGAACGATCTGTGAGATCACTGATCCGCGCCCAACCAACAGCGCT 963
 QY 793 AGCTGTTCTGCAACATGTCATGAGCGCAAGCTTGAATCTGGAAGAAAGTGGCC 852
 DB 964 AAGAGCATCAATCGGCGCGCGCGCTTCTACGCGCACCGGCGACATCATCGCGCAG 1023
 QY 853 GTGCAATCGCGCGCGCTGAGTCAAGCGGCGCGCGCTGCAAGGACATCGCGCATC 912
 DB 1024 ATTCGCGAGGCGCACTGCAACATCAACGCAAGGCGCAATGCAACCAACCTTGAAGATC 1083
 QY 913 GAGCTCAACGAGATGATCGCGAGTGGACCTGCTGAAGTACCGCTCTTGGCTATGAGC 972
 DB 1084 G---TGGAGAACTGCGCGAGAGTTCGGCAACAAGAACATCATCTTTCAACAGCAGC 1140
 QY 973 CACTCTTGGCGGTGTGCTGCACTACAGGTGCGAGGCGCGCGTGAAGCTGCGCGAG 1032
 DB 1141 AGCGGCGGCGACCGCGAGATCGTGTTCACAGCTTCACTGCGGCGCGCGGTTCTTCAAC 1200
 QY 1033 GACATGCAACCGAGTGAAGCGCGGATGAGTGTCTCAATGAGCGCATGAGTGTGCTG 1092
 DB 1201 TGCACACACGAGCTGTTCAACAGCACTGGAATATCAACGAGAGGTGAAACAAGACC 1260
 QY 1093 CGGAGAGGATGCGCGGCGCGGCTATGCGAGACGACATCTGATCGTGGGAGAG 1152
 DB 1261 AAGGAGAAAGACATCATCTGCTGCTGCGCATTCGCGCATCATCAATGTGCGAG 1320
 QY 1153 GACGCTGCGGAAACATCAACGCGCTTTCGTTTCGATTCGGAACAACAATCATTCGCAA 1211
 DB 1321 GAGGTGGGCAAGGACATGATACGCGCGCGCGCGCATTCGCGGCGAGTCAAGTGAAGCA 1379

RESULT 14
 US-09-475-515-58
 ; Sequence 58, Application US/09475515A
 ; Patent No. 6602705
 ; GENERAL INFORMATION:
 ; APPLICANT: BARNETT, Susan
 ; APPLICANT: ZUR MEGEDE, Jan
 ; APPLICANT: SRIVASTAVA, Indresh
 ; APPLICANT: LIAN, Ying
 ; APPLICANT: HARTOG, Karin
 ; APPLICANT: LIU, Hong
 ; APPLICANT: GREER, Catherine
 ; APPLICANT: SELBY, Mark
 ; APPLICANT: WALKER, Christopher
 ; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
 ; FILE REFERENCE: 1621.002
 ; CURRENT APPLICATION NUMBER: US/09/475,515A
 ; CURRENT FILING DATE: 1999-12-30
 ; NUMBER OF SEQ ID NOS: 90
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 58
 ; LENGTH: 2181
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: gp140TM.modUS4
 US-09-475-515-58

Query Match 5.0%; Score 60.6; DB 4; Length 2181;
 Best Local Similarity 42.6%; Pred. No. 0.0015;
 Matches 434; Conservative 0; Mismatches 579; Indels 6; Gaps 2;

QY 196 CGCAAGTACGGCATGTGATGACCAACCAACGCCAGATTTGGCCGGCATGAC 255
 DB 364 CCCCTGTGCTGACCCCTGAACTGACCGACAGTGAACCGGACCAACCGGACCAAC 423
 QY 256 GGGGGCCAGCCCTGGGCGCCGAGTTTGGGCAACAATCACTACACCGTGGCCGC 315
 DB 424 AGACACCGCGCAACCAACGACCAAGGCGGACCAACGACCAACGACGACGAC 483
 QY 316 GACAAATTTATGCGCGCGCGCGAGCT---GACCAAGGCGCGCAAGCGCATCGGATC 372
 DB 484 AGCTGGAGAGAGATGCCCGAGGGCGAGATCAAGAACTGCAAGCTTCAATCAACCAACGAC 543
 QY 433 GAGTTGTGACATCAAGCCGCTGATGTGATGCGGACCAATCAAGTCTGGAAGAG 492
 DB 604 GACAAAGCAACCGCAGCTACCGCTGATCAACGCAACCAAGCGTATCAACCCAGGCC 663
 QY 493 CAGAACTGATCGCGAAGGCGCGCGCTGTGTGAGTGGCGGCGCGCGCTGCGGCGCT 552
 DB 664 TGCCCCAAGGTGAGCTTCAAGCCATCCCATCACTACCTGCGCCCGCGCTTGGCC 723
 QY 553 GCCATCAAGGCGCGGCTGCGCGAGATGAAAGTGGCATGCGCACCAACCAATGCGATGATC 612
 DB 724 ATCTGAAAGTGAAGGCAAGAAAGTTCAACGCAACGCGCCCTGCAAGAAAGTGAAGAC 783
 QY 613 CGGGAATGCGCAATCTTCCCTTGTGTGAGTGAAGTGAACCTGGAACCTGTTCCAG 672
 DB 784 GTGCAATGCAACCGAGATCGGCGCGCGCTTCTACGCGCACCGGCGACATCATCGCGCAG 843
 QY 673 TCGGGCATCAACCGAGCGCGGCAATTCGGTCAACCAAGCGCATGCGCATCGCG 732
 DB 844 CTGGCGAGAGAGATGCTGTGCTGCTGCGAACTTCAACGCAACGCGCAAGACATC 903
 QY 733 GACATCTTCTGCTCAACACTTCCGAGATCTTGGGCTACTACACCGCGCTGAGAGCG 792
 DB 904 ATCTGACACTGAACGATCTGTGAGATCACTGATCCGCGCCCAACCAACAGCGCT 963

QY 793 AGCGTGTTCGAGCATTCATGACGCGCAGCTTCGACATCTGGAGAAAGACTGGCC 852
 DB 964 AAGAGCATTCACATCGGCCCCCGCGCTTCTACGCCACCGGCACATCATCGCCGAC 1023
 QY 853 GTGCATTCGCGCGGGGTCTGAGCTGATCAAGCGGGGCGCGGTGCAAGAGCATTCGCATC 912
 DB 1024 ATCCGCGAGGCCCATCTGCAACATCAGCAAGGCCAATTGACCAACACCTTCAGACAGATC 1083
 QY 913 GAGCTCAACGAGATGTACCGCGAGTGGGACTGTGATTAACCGCTCTTTCGGCTATGAC 972
 DB 1084 G---TGAGAAAGCTGGCGACAGAGATTGGGCAACAAAGACATCATCTTCAACAGCAGC 1140
 QY 973 CACTCTTCGCGGTGTGTGCACTACAGCTGCGAGGCGCGGTGAGAGCTGGCGGAG 1032
 DB 1141 AGCGCGCGGACCCCGAGATCTGTTCACAGCTTCACTGCGCGCGGCAAGTCTTCTAC 1200
 QY 1033 GACATCGAACCGAGTGAAGCCCGGCAATGTGTCTCATGTAGAGCCGATGTGATGTCG 1092
 DB 1201 TGCACACACGACGCTGTTCACAGCAGCTTGAAATATCACAGAGAGGTGAACAGACC 1260
 QY 1093 CCGAGGGGATCCCGGATCCGCGGCTATTCGAGACAGACATCTCTGATCGGGAG 1152
 DB 1261 AAGGAAAGACACATCATCTCTGCGCTGCGGATCGCAGATCATCAATGTGGAG 1320
 QY 1153 GACGCTCCGAGAACATCACCGGCTTCCGTTGCGTCCGGAACAACATCATCCGCA 1211
 DB 1321 GAGGTGGGCAAGGCCATGTACGCCCCCCCCCATCCGCGGCAAGTGCAGCAGCAA 1379

RESULT 15
 US-09-475-515-64
 Sequence 64, Application US/09475515A

GENERAL INFORMATION:
 Patent No. 6602705
 APPLICANT: BARNETT, Susan
 APPLICANT: ZUR MEGEDE, Jan
 APPLICANT: SRIVASTAVA, Indresh
 APPLICANT: LIAN, Ying
 APPLICANT: HARTOG, Karin
 APPLICANT: LIU, Hong
 APPLICANT: GREER, Catherine
 APPLICANT: SELBY, Mark
 APPLICANT: WALKER, Christopher
 TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
 FILE REFERENCE: 1621.002
 CURRENT APPLICATION NUMBER: US/09/475,515A
 CURRENT FILING DATE: 1999-12-30
 NUMBER OF SEQ ID NOS: 90
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 64
 LENGTH: 2634
 TYPE: DNA
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Description of Artificial Sequence: gp160.modus4
 US-09-475-515-64

Query Match 5.0%; Score 60.6; DB 4; Length 2634;
 Best Local Similarity 42.6%; Pred. No. 0.0015;
 Matches 434; Conservative 0; Mismatches 579; Indels 6; Gaps 2;

QY 196 CCGAAGTACGCGATGTCTATCGACCAACAGCCACGATTTCCGCGCGCATCGAC 255
 DB 364 CCCCTGTGTGATCCTTGAACTGACGCAAGCTAGTCGCGGACCAACGCGCACAC 423
 QY 256 GCGCGGCAAGCTTCGCGGCGGAGCTTTCGGGACACATCATCTACCGAGTGGCGCG 315
 DB 424 AGCACCAGCGGACCAACAGACCAAGCGGACCAACAGCAGCAGCAACAGCAGCAG 483
 QY 316 GACATTTTATTCGCGCGGTGCGCGAGCT---GACCAAGGCGGCAAGGCGCATCGGCATC 372

DB 484 AGCTGGAGAGATGCCCCGAGGCGAGATCAAGAACTGACGTTCAACATCACACAGC 543
 QY 373 GAGTTGACACAGTAAATCTGACTTCCCGCGCCAGCTTCAGAGAAAGCCCTACCGGCGTC 432
 DB 544 GTGCGCGCAAGAGTGCACAAAGAGAGTACAGCTGTCTTCAAGCTGAGAGTGTGCCATC 603
 QY 433 GAGTTCTGTGACATGACGACGCTCTGATGTGATATGCGCATCATCAATCGCTCGAAG 492
 DB 604 GACAAAGCAAGCCAGCTTACCGCTGTATCACTGCAACACCAAGGTATACACCAAGGCC 663
 QY 493 CAGAACTGATTCGCGAAGGCGCGCGTGTGTGACGTGCGGCGGCGCGCTGCGGCT 552
 DB 664 TGCCCAAGAGTGAAGTTTGAAGCCATCCCATTCACATCTGAGCCCGCGGCTTCCGC 723
 QY 553 GGCATGAAGCGCGGCTGCCGAGCATGAATGGAGATTCGCCACCAACCAATGCCATGATC 612
 DB 724 ATCTGAAGTGCAGGACAAAGAGTTCAACGCGACCGGCGCTGCAAGAAAGTGAAGACC 783
 QY 613 CGCGAGATCGCAATGTTCCCTTGTGGAGCTGATGACACTGGAAGCTGTTCAG 672
 DB 784 GTGCAATGACATCCAGCATTCGCGCTGTGTGACACCAAGCTGTCTGAAACGCGAGC 843
 QY 673 TCGGCGATCAACACCGAGGCGCGCAATCCGATCAACCAACGATGTGATTCGCG 732
 DB 844 CTGCGCGAGAGAGATGTGTGTGCTCCGAGAACTTCAACGACAGCCAAAGACATC 903
 QY 733 GACATCTTTGTGCTAACAACCTTCCGATGATCTTCCGCTACTAACCAGCGCTGAGAGCG 792
 DB 904 ATCTGAGCTGTGAAGATCCGTGTGAGATCACTGACATCCGCCCAACAAACACGCT 963
 QY 793 AGCGTGTTCGAGCATGTGATGACGCGACGCTTCGACATCTGGAGAAAGACGTGGCC 852
 DB 964 AAGAGCATTCACATGGCGCCCGCGGCTTCTACGCCACCGGAGATCATTCGGCGAC 1023
 QY 853 GTGCATTCGCGGGCTGAGCTGATCAAGCGCGCGCGCTGCAAGAGCATTCGCCATC 912
 DB 1024 ATCCGCGAGGCCCATCTGCAACATCAGCAAGGCCAATCGAACCAACCTTCAGAGATC 1083
 QY 913 GAGCTCAACGAGATGTACCGGAGTGGGACTCTGTAAGTACCGCTCTTCCGCTATGGC 972
 DB 1084 G---TGAGAAAGCTGGCGAGCAGTTCGCAACAAAGAACCATCATCTTCAACAGCAGC 1140
 QY 973 CACTCTTCGCGGTGTGACACTACTACGCTGCGAGGCGCGGTGAGAGCTGGCGGAG 1032
 DB 1141 AGCGCGCGGACCCGAGATGTGTTCACAGCTTCAACTGGCGGCGAGTTCTTCTAC 1200
 QY 1033 GACATCGAACCGAGCTGAAGCCCGCATGTGTCTCATGAGCCGATGTGATGTCG 1092
 DB 1201 TGCACACACGACGCTGTTCACAGCAGCTTGGAATATCACAGAGAGGTGAACAGACC 1260
 QY 1093 CCGAGGGGATCCCGGATCCGCGGCTATTCGCGACAGACATCTCTGATCTCGGGAG 1152
 DB 1261 AAGGAAAGACACCATCATCTCTGCGGATTCGCGCATTCGCAATCATCAAGTGGCAG 1320
 QY 1153 GACGCTCCGAGAACATCACCGGCTTCCGTTGCGTCCGGAACAACATCATCCGCA 1211
 DB 1321 GAGGTGGGCAAGGCCATGTACGCCCCCCCCCATCCGCGGCAAGTGCAGCAGCAA 1379

Search completed: July 7, 2005, 23:34:22
 Job time : 240 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comphen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 7, 2005, 22:16:36 ; Search time 828 Seconds

(without alignments)
9189.996 Million cell updates/sec

Title: US-10-807-228a-2

Perfect score: 1212
Sequence: 1 ATGACTGACGACATGTGCA.....AACACACATCATCCGCAAC 1212

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 6330943 seqs, 3139157217 residues

Total number of hits satisfying chosen parameters: 12661886

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_NA.*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:*
22: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:*
24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	984.8	81.3	1215	US-10-251-078-1	Sequence 17, Appli
2	984.8	81.3	1215	US-10-251-078-17	Sequence 17, Appli
3	981.6	81.0	1215	US-10-251-078-9	Sequence 9, Appli
4	980	80.9	1212	US-10-251-078-11	Sequence 11, Appli
5	980	80.9	1212	US-10-251-078-13	Sequence 13, Appli
6	980	80.9	1215	US-10-251-078-15	Sequence 15, Appli
7	978.4	80.7	1212	US-10-251-078-21	Sequence 21, Appli

8	976.8	80.6	1212	US-10-251-078-19	Sequence 19, Appli
9	73.8	6.1	2733	US-10-437-963-88003	Sequence 88003, A
10	71.8	5.9	1262	US-10-428-115-109294	Sequence 109294, A
11	63.8	5.8	1483	US-10-437-963-38426	Sequence 38426, A
12	69.6	5.7	1212	US-10-411-910A-222	Sequence 222, App
13	68.6	5.7	1169	US-10-437-963-57011	Sequence 57011, A
14	68	5.6	834	US-10-156-761-4932	Sequence 4932, Ap
15	68	5.6	985	US-10-363-345A-31407	Sequence 31407, A
16	68	5.6	985	US-10-363-345A-31408	Sequence 31408, A
17	68	5.6	985	US-10-363-345A-31407	Sequence 31407, A
18	68	5.6	985	US-10-363-345A-31408	Sequence 31408, A
19	68	5.6	9025608	US-10-156-761-1	Sequence 1, Appli
20	67.8	5.6	1227	US-10-399-693-15	Sequence 15, Appli
21	67.6	5.6	1566	US-10-817-483-10	Sequence 10, Appli
22	67.2	5.5	65140	US-10-203-295-1	Sequence 1, Appli
23	67.2	5.5	125401	US-10-203-295-35	Sequence 35, Appli
24	66.6	5.5	2052	US-10-282-122A-11731	Sequence 11731, A
25	66	5.4	2960	US-10-437-963-34365	Sequence 34365, A
26	65.6	5.4	1516	US-10-437-963-29365	Sequence 29365, A
27	65.4	5.4	2569	US-10-437-963-88004	Sequence 88004, A
28	65.4	5.4	4512	US-10-156-761-2222	Sequence 2222, Ap
29	65.4	5.4	138203	US-10-819-386A-1	Sequence 1, Appli
30	65.4	5.4	9025608	US-10-156-761-1	Sequence 1, Appli
31	65.2	5.4	2712	US-10-282-122A-24663	Sequence 24663, A
32	65	5.4	573	US-10-156-761-3931	Sequence 3931, A
33	64.6	5.3	9546	US-10-463-074-2	Sequence 2, Appli
34	64.4	5.3	2037	US-10-228-063-48	Sequence 48, Appli
35	64	5.3	1053	US-10-156-761-5241	Sequence 5241, Ap
36	64	5.3	1294	US-09-748-033-2	Sequence 2, Appli
37	63.6	5.2	690	US-10-006-922-27	Sequence 27, Appli
38	63.6	5.2	1161	US-10-081-664-19	Sequence 19, Appli
39	63.4	5.2	1161	US-10-282-122A-11406	Sequence 11406, A
40	63.4	5.2	2026	US-10-437-963-32278	Sequence 32278, A
41	63.4	5.2	2649	US-10-437-963-112	Sequence 112, App
42	63.2	5.2	9785	US-10-359-120-176	Sequence 176, App
43	63	5.2	1752	US-10-411-910A-234	Sequence 234, App
44	63	5.2	14335	US-10-156-761-412	Sequence 412, App
45	63	5.2	100000	US-10-156-761-15103	Sequence 15103, A

ALIGNMENTS

RESULT 1
US-10-251-078-1
; Sequence 1, Application US/10251078
; Publication No. US20030119084A1
; GENERAL INFORMATION:
; APPLICANT: Roche Diagnostics GmbH
; APPLICANT: Shao, Zhixin
; APPLICANT: Schmuck, Rainer
; APPLICANT: Kratzsch, Peter
; APPLICANT: Kenkles, Janet
; APPLICANT: Weisner, Harald
; TITLE OF INVENTION: Variants of an Erwinia-type creatinase
; FILE REFERENCE: 20981 EP
; CURRENT APPLICATION NUMBER: US/10/251.078
; CURRENT FILING DATE: 2002-09-20
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1215
; TYPE: DNA
; ORGANISM: Erwinia sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1215)
US-10-251-078-1

Query Match 81.3%; Score 984.8; DB 15; Length 1215;
Best Local Similarity 88.3%; Pred. No. 1.2e+258;
Matches 1070; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

QY 1 ATGACTGAGCAATGTTGACGATGTAATGAGCAACGCGGAGAAAGATTATTCGCGG 60
DB 1 ATGACTGAGCAATGTTGACGATGTAATGAGCAACGCGGAGAAAGATTATTCGCGG 60
QY 61 TTTTGGATGCGAGATGACCGCGCCAAAGCAAGTTTCGCGGCTGATGCGCAAGAC 120
DB 61 TTTTCCGATGCGGAGATGACCGCGCCAGAGATGCGGCGCTGATGCGCAAGAC 120
QY 121 AATGTCATGCGGCGCTGTTCACTCTTATCACTGATCACTACTATTCGCGCTGCTG 180
DB 121 GACGTGACGCTGCGCTGTTCACTCTTATCACTGATCACTACTATTCGCTGCTG 180
QY 181 TACTGCTATTTGCGACGCAAGTACGCGCATGTCATGACCAAGCAAGCGCAAGATT 240
DB 181 TACTGCTATTTGCGCGCAAAATACGCGCATGTCATGACCAAGCAAGCGCAAGATT 240
QY 241 TCGGCGCGCATGACCG 300
DB 241 TCGGCGCGCATGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
QY 301 ACCGACTGCGCGCGCGCAATTTCTATGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
DB 301 ACGGACTGCGCGCGCGCAACTTTCTACGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
QY 361 CGCATGCGCATGAGTTGACACCGCTCATCTTCGCTCGCGCGCGCGCGCGCGCGCG 420
DB 361 CGCATGCGCATGAGTTGACACCGCTCATCTTCGCTCGCGCGCGCGCGCGCGCGCG 420
QY 421 CTACGCGCGCGCTGATGCTGTCGATCAACCGCGCGCGCGCGCGCGCGCGCGCGCG 480
DB 421 CTGCGCGCGCGCTGATGCTGTCGATCAACCGCTGATGCTGATGCGCGCGCGCGCG 480
QY 481 TCGCTGAG 540
DB 481 TCGCTGAG 540
QY 541 GCGTGGGCGCGCGCATCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
DB 541 GCGTGGGCGCGCGCATCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
QY 601 AATGCGATGATCGCGAGATGCGCAATGCTTCCCTTCTGTCGAGCTGATGAGACCTTG 660
DB 601 AATGCGATGATCGCGAGATGCGCAATGCTTCCCTTCTGTCGAGCTGATGAGACCTTG 660
QY 661 ACTGCTTCACTGCGCGCATCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
DB 661 ACTGCTTCACTGCGCGCATCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
QY 721 GTGCAATCGCGGAGCATCTCTTCTGTCACCTTCCGATGATCTTCCGCTACTACAC 780
DB 721 GTGCAATCGCGGAGCATCTCTTCTGTCACCTTCCGATGATCTTCCGCTACTACAC 780
QY 781 GCGCTGAGAGCGACGCTGTTCTGCGACCAATGTCATGACGCGCGCTGCACTTGGAG 840
DB 781 GCGCTGAGAGCGACGCTGTTCTGCGACCAATGTCATGACGCGCGCTGCACTTGGAG 840
QY 841 AAGAAAGTGGCGCTGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
DB 841 AAGAAAGTGGCGCTGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
QY 901 GACATGCGCATGAGCTTCAACGAGATGTAACCGGAGATGAGATGCTGTAAGATACCGCTCC 960
DB 901 GATATGCGCATGAGCTTCAACGAGATGTAACCGGAGATGAGATGCTGTAAGATACCGCTCC 960
QY 961 TTCGCGCTATGCGCACTCTTCCGCGCTGCTGCTCACTACGCGCGCGCGCGCGCG 1020
DB 961 TTCGCGCTATGCGCACTCTTCCGCGCTGCTGCTCACTACGCGCGCGCGCGCGCG 1020
QY 1021 GAGCTGCGGAGAGCATGACACCGAGCTGAGAGCGCGCGCGCGCGCGCGCGCGCG 1080
DB 1021 GAACTGCGGAGAGCATGACACCGAGCTGAGAGCGCGCGCGCGCGCGCGCGCGCG 1080
QY 1081 ATGCTGATGCTGCGGAGGAGCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1140

DB 1081 ATGCTGATGCTGCGGAGGAGCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1140
QY 1141 ATGCTGCGGAGAGAGAGCGCTGCGGAGAGCATCAACCGCGCTTCCGCTGCTGCGAACAAC 1200
DB 1141 ATGCTGCGGAGAGAGAGCGCGCGGAGAGCATTAACCGGATTCCTTCCGCTGCGAACAAC 1200
QY 1201 ATCATCGCGCAAC 1212
DB 1201 ATCATCGCGCAAC 1212

RESULT 2
US-10-251-078-17
; Sequence 17, Application US/10251078
; Publication No. US20030119084A1
; GENERAL INFORMATION:
; APPLICANT: Roche Diagnostics GmbH
; APPLICANT: Shao, Zhixin
; APPLICANT: Schmuck, Rainer
; APPLICANT: Kratzsch, Peter
; APPLICANT: Kenkies, Janet
; APPLICANT: Welser, Harald
; TITLE OF INVENTION: Variants of an Erwinia-type creatinase
; FILE REFERENCE: 20981 EP
; CURRENT APPLICATION NUMBER: US/10/251,078
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 1215
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:variant CTGc2
; NAME/KEY: CDS
; LOCATION: (1)..(1215)
US-10-251-078-17

Query Match 81.3%; Score 984.8; DB 15; Length 1215;
Best Local Similarity 88.3%; Pred. No. 1.2e-258;
Matches 1070; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

QY 1 ATGACTGAGCAATGTTGACGATGTAATGAGCAACGCGGAGAAAGATTATTCGCGG 60
DB 1 ATGACTGAGCAATGTTGACGATGTAATGAGCAACGCGGAGAAAGATTATTCGCGG 60
QY 61 TTTTGGATGCGGAGATGACCGCGCCAAAGCAAGTTTCGCGGCTGATGCGCAAGAC 120
DB 61 TTTTCCGATGCGGAGATGACCGCGCCAGAGATGCGGCGCTGATGCGCAAGAC 120
QY 121 AATGTCATGCGGCGCTGTTCACTCTTATCACTGATCACTACTATTCGCGCTGCTG 180
DB 121 GACGTGACGCTGCGCTGTTCACTCTTATCACTGATCACTACTATTCGCTGCTG 180
QY 181 TACTGCTATTTGCGACGCAAGTACGCGCATGTCATGACCAAGCAAGCGCAAGATT 240
DB 181 TACTGCTATTTGCGCGCAAAATACGCGCATGTCATGACCAAGCAAGCGCAAGATT 240
QY 241 TCGGCGCGCATGACCG 300
DB 241 TCGGCGCGCATGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
QY 301 ACCGACTGCGCGCGCGCAATTTCTATGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
DB 301 ACGGACTGCGCGCGCGCAACTTTCTACGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
QY 361 CGCATGCGCATGAGTTGACACCGCTCATCTTCGCTCGCGCGCGCGCGCGCGCGCG 420
DB 361 CGCATGCGCATGAGTTGACACCGCTCATCTTCGCTCGCGCGCGCGCGCGCGCGCG 420
QY 421 CTACGCGCGCGCTGATGCTGTCGATCAACCGCGCGCGCGCGCGCGCGCGCGCGCG 480

```
Db 421 CTGCCGCGGTGAGTTCCTGATATCGGTCAACCGTTCATGTGATGCGACACGGTCAAG 480
Qy 481 TGGCTCGAAGACAGAAAGCTGATCCGCCGAAGGCGCCGCTGTGTATGCTCGCGCGCG 540
Db 481 TGGCTCGAAGACAGAAAGCTGATCCGCCGAAGGCGCCGCTGTGTATGCTCGCGCGCG 540
Qy 541 GCCTGCGCGGCTGCCATCAAGGCGCGGCTGCCGAGCATGATGCGCATCGCAACACC 600
Db 541 GCCTGCGGTCGCGCGCTCAAGGCGCGGCTGCCGAGCATGATGCGCATCGCAACACC 600
Qy 601 AATGCGATGATCCGCGAGATCCGCAATGCTTCCCTTCGTGAGTGTGATGACACTCG 660
Db 601 AATGCGATGATCCGCGAGATCCGCAAGTGTTCCTTCGTGAGTGTGATGACACTCG 660
Qy 661 ACCTGCTTCATCGCGGCTCAACACCGAGCGCGCAATCCGCTCAACCGCATC 720
Db 661 ACCTGCTTCATCGCGGCTCAACACCGAGCGCGCAATCCGCTCAACCGCATC 720
Qy 721 GTGCAATCCGCGGCAATCCTTTCGCTCAACCTTCGCAATGATCTTCGCTCAATCACC 780
Db 721 GTGCAATCCGCGGCAATCCTTTCGCTCAACCTTCGCAATGATCTTCGCTCAATCACC 780
Qy 781 GCGCTGAGAGCGACGCTGTTCTGCGACATGTGATGACGCGAGCTCGACATCTGGAG 840
Db 781 GCGCTGAGAGCGACGCTGTTCTGCGACATGTGATGACGCGAGCTCGACATCTGGAG 840
Qy 841 AAGAACTGTCGCTGATCGCGCGCGGCTCGAGCTGATCAAGCCGCGCGCGCTGCAAG 900
Db 841 AAGAACTGTCGCTGATCGCGCGCGGCTCGAGCTGATCAAGCCGCGCGCGCTGCAAG 900
Qy 901 GACATCGCATGATGATCAAGAGATGTAACGCGAGTGGAGCTGTGAAATACCGCTCC 960
Db 901 GATATCGCATGATGATCAAGAGATGTAACGCGAGTGGAGCTGTGAAATACCGCTCC 960
Qy 961 TTCCGCTATGCGCACTCTTTCGCGCTGTGCTGCTCACTACGCTCGACGCGCGCTG 1020
Db 961 TTCCGCTATGCGCACTCTTTCGCGCTGTGCTGCTCACTACGCTCGACGCGCGCTG 1020
Qy 1021 GAGCTCGCGAGGACATGACACCGAGCTGAAAGCCCGGCTGATGTGTCTCCATGAGCCG 1080
Db 1021 GAGCTCGCGAGGACATGACACCGAGCTGAAAGCCCGGCTGATGTGTCTCCATGAGCCG 1080
Qy 1081 ATGCGATGCTGCGCGGAGGATGCGCGCTGCGCGGCTATCGGAGACACATCTCG 1140
Db 1081 ATGCGATGCTGCGCGGAGGATGCGCGCTGCGCGGCTATCGGAGACACATCTCG 1140
Qy 1141 ATCGTCGCGGAGGACGCTGCGCGAGACATCAACCGCTTCGCGTTCGCGGACACAC 1200
Db 1141 ATCGTCGCGGAGGACGCTGCGCGAGGACATCAACCGCTTCGCGTTCGCGGACACAC 1200
Qy 1201 ATCATCCGCAAC 1212
Db 1201 ATCATCCGCAAC 1212
```

```
RESULT 3
US-10-251-078-9
; Sequence 9, Application US/10251078
; Publication No. US20030119084A1
; GENERAL INFORMATION:
; APPLICANT: Roche Diagnostics GmbH
; APPLICANT: Shao, Zhixin
; APPLICANT: Schumuck, Rainer
; APPLICANT: Kretzsch, Peter
; APPLICANT: Kenkies, Janet
; APPLICANT: Weisner, Harald
; TITLE OF INVENTION: Variants of an Erwinia-type creatinase
; FILE REFERENCE: 20981 EP
; CURRENT APPLICATION NUMBER: US/10/251.078
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn version 3.1
```

```
; SEQ ID NO 9
; LENGTH: 1215
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: variant CTIM24
; NAME/KEY: CDS
; LOCATION: (1)..(1215)
US-10-251-078-9

Query Match      81.0%; Score 981.6; DB 15; Length 1215;
Best Local Similarity 88.1%; Pred. No. 9e-258;
Matches 1068; Conservative 0; Mismatches 144; Indels 0; Gaps 0;

Qy 1 ATGACTGACGACATGTTGACGATGATGAAATGACCAACGCGAGAAAGATTATTCGCCG 60
Db 1 ATGACTGACGACATGTTGACGATGATGAAATGACCAATGCGAGAAAGATTATTCGCCG 60
Qy 61 TTTTCGATGCGCGAGATGACGCGCGCGCAAGGTGCGGCGCTGAGTGGCCGAAC 120
Db 61 TTTTCGATGCGCGAGATGACGCGCGCGCAAGGTGCGGCGCTGAGTGGCCGAAC 120
Qy 121 AATGTCGATGCGCGCTGTTCACTTATCACTGATCACTAATCTAATCCGCTGCTG 180
Db 121 AATGTCGATGCGCGCTGTTCACTTATCACTGATCACTAATCTAATCCGCTGCTG 180
Qy 121 GACGTGACGCTGCGCTGTTCACTTATCACTGATCACTAATCTAATCCGCTGCTG 180
Db 121 GACGTGACGCTGCGCTGTTCACTTATCACTGATCACTAATCTAATCCGCTGCTG 180
Qy 181 TACTGCTATTTTGGACGCAAGTACGCGCATGTGATGACCAACAGCCACGACGATT 240
Db 181 TACTGCTATTTTGGACGCAAGTACGCGCATGTGATGACCAAGCCACGACGATT 240
Qy 241 TCGCGCGGATCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
Db 241 TCGCGCGGATCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
Qy 301 ACCGATGCGCGCGCGCGCAATTTCTATGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
Db 301 ACGATGCGCGCGCGCGCGCAATTTCTATGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
Qy 361 GCGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
Db 361 GCGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
Qy 421 CTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
Db 421 CTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
Qy 481 TCGCTCGAAGACAGAAAGCTGATCCGCGAGGATGCGCGGCTATCTGCGAGCTGCGG 540
Db 481 TCGCTCGAAGACAGAAAGCTGATCCGCGAGGATGCGCGGCTATCTGCGAGCTGCGG 540
Qy 541 GCCTGCGCGGCTGCGCATCAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
Db 541 GCCTGCGCGGCTGCGCATCAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
Qy 601 AATGCGATGATCCGCGAGATGCGCAATGCTTCCCTTCGTGAGTGTGATGACACTCG 660
Db 601 AATGCGATGATCCGCGAGATGCGCAATGCTTCCCTTCGTGAGTGTGATGACACTCG 660
Qy 661 ACCTGCTTCATCGCGGCTCAACACCGAGCGCGCAATCCGCTCAACCGCATC 720
Db 661 ACCTGCTTCATCGCGGCTCAACACCGAGCGCGCAATCCGCTCAACCGCATC 720
Qy 721 GTGCAATCCGCGGCAATCCTTTCGCTCAACCTTCGCAATGATCTTCGCTCAATCACC 780
Db 721 GTGCAATCCGCGGCAATCCTTTCGCTCAACCTTCGCAATGATCTTCGCTCAATCACC 780
Qy 781 GCGCTGAGAGCGACGCTGTTCTGCGACATGTGATGACGCGAGCTCGACATCTGGAG 840
Db 781 GCGCTGAGAGCGACGCTGTTCTGCGACATGTGATGACGCGAGCTCGACATCTGGAG 840
Qy 841 AAGAACTGTCGCTGATCGCGCGCGGCTCGAGCTGATCAAGCCGCGCGCGCTGCAAG 900
Db 841 AAGAACTGTCGCTGATCGCGCGCGGCTCGAGCTGATCAAGCCGCGCGCGCTGCAAG 900
```


Db 841 AAGAACTGCGCGTGCACCGCGCGCTCGAATCATCAAGCCGAGTGCCTGCAAG 900
Qy 901 GACATCGCCATGAGCTCAACAGATGTACCGCGAGTGGACCTGCTGAATACCGCTCC 960
Db 901 GATATCGCATGAACTCAACAGATGTACCGCGAGTGGACTCTGTAAGTACCGCTCC 960
Qy 961 TTGGCTATGAGCACTCTTGGCGGTGCTGTGCCACTACTACGTGCGAGGCGCGCTG 1020
Db 961 TTGGCTATGAGCACTCTTGGCGGTGCTGTGCCACTACTACGTGCGAGGCGCGCTG 1020
Qy 1021 GAGCTGCGCGAGGATGACACCGAGCTGAGCCCGGATGCTGTCTCCATGAGCCG 1080
Db 1021 GAACTGCGCGAGGATGACACCGAGCTGAGCCCGGATGCTGTCTCCATGAGCCG 1080
Qy 1081 ATGCTGATGCTGCGCGAGGCAATGCGCGGTGCGCGGCTATCGAGACACATCTTG 1140
Db 1081 ATGCTGATGCTGCGCGAGGCAATGCGCGGTGCGCGGCTATCGAGACACATCTTG 1140
Qy 1141 ATCTGCGGAGAGAGAGCTGCGAGACATCAACCGCTTCCGTTGCTCGGACACAC 1200
Db 1141 ATCTGCGGAGAGAGAGCTGCGAGACATCAACCGCTTCCGTTGCTCGGACACAC 1200
Qy 1201 ATCATCCGCAAC 1212
Db 1201 ATCATCCGCAAC 1212

RESULT 4

US-10-251-078-11

/ Sequence 11, Application US/10251078
/ Publication No. US20030119084A1
/ GENERAL INFORMATION:
/ APPLICANT: Roche Diagnostics GmbH
/ APPLICANT: Shao, Zhixin
/ APPLICANT: Schmuck, Rainer
/ APPLICANT: Kratzsch, Peter
/ APPLICANT: Kenkies, Janet
/ APPLICANT: Weisner, Harald
/ TITLE OF INVENTION: Variants of an Erwinia-type creatinase
/ FILE REFERENCE: 20981 EP
/ CURRENT APPLICATION NUMBER: US/10/251,078
/ CURRENT FILING DATE: 2002-09-20
/ NUMBER OF SEQ ID NOS: 32
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 11
/ LENGTH: 1212
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: variant CT2m9
/ NAME/KEY: CDS
/ LOCATION: (1)..(1212)
US-10-251-078-11

Query Match 80.9%; Score 980; DB 15; Length 1212;

Best Local Similarity 88.0%; Pred. No. 2.5e-257;

Matches 1067; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

Qy 1 ATGACTGACGACATGTTGACGCTGATGAATAAGCACACGAGGAGAAAATATTTCGCCG 60
Db 1 ATGACTGACGACATGTTGACGCTGATGAATAAGCACACGAGGAGAAAATATTTCGCCG 60
Qy 61 TTTTCGATGCGGAGTGAACCGCGCCAAAAGACGTTGCGGCTGATGAGCAAGAAC 120
Db 61 TTTTCGATGCGGAGTGAACCGCGCCAAAAGACGTTGCGGCTGATGAGCAAGAAC 120
Qy 121 AATGCTGATGCGGCTGTTTCACTCTTATCACTGACATCACTATTTCGGCTGGCTG 180
Db 121 AATGCTGATGCGGCTGTTTCACTCTTATCACTGACATCACTATTTCGGCTGGCTG 180
Qy 181 TACTGCTATTTGGACGCAAGTACGCAATGCTCATGACCAACAACGCCACGACGATT 240
Db 181 TACTGCTATTTGGACGCAAGTACGCAATGCTCATGACCAACAACGCCACGACGATT 240

Db 181 TACTGCTATTTGGACGCAAGTACGCAATGCTCATGACCAACAACGCCACGACGATT 240
Qy 241 TCGGCGGACATGACCGCGGCAAGCTTGGCGCGCAAGCTTTCGAGCAACATCACTTAC 300
Db 241 TCGGCGGACATGACCGCGGCAAGCTTGGCGCGCAAGCTTTCGAGCAACATCACTTAT 300
Qy 301 ACCGACTGCGCGCGGACAAATTTCTATGCGCGCTGTGCGGCACTGACCAACGCGCC 360
Db 301 ACCGACTGCGCGCGGACAAATTTCTATGCGCGCTGTGCGGCACTGACCAACGCGCC 360
Qy 361 CGCATCGGATCGAGTTGACCAAGCTGCAATCTGCACTTCGCGCGCACTGCGAGAGCC 420
Db 361 CGCATCGGATCGAGTTGACCAAGCTGCAATCTGCACTTCGCGCGCACTGCGAGAGCC 420
Qy 421 CTACCGGCGCTGAGTTGTCGACATCAACGCAAGCTTCGATGTGATGCGCACTCAAG 480
Db 421 CTACCGGCGCTGAGTTGTCGACATCAACGCAAGCTTCGATGTGATGCGCACTCAAG 480
Qy 481 TCGCTGAAAGACGAACTGATCCGCGAAGCGCGCGCGGTGTGATGAGTGGCGCGCG 540
Db 481 TCGCTGAAAGACGAACTGATCCGCGAAGCGCGCGCGGTGTGATGAGTGGCGCGCG 540
Qy 541 GCTGCGCGGCTGCGCATCAAGCGCGCGTGCAGCATGAAGTGGCGATGCGCAAC 600
Db 541 GCTGCGCGGCTGCGCATCAAGCGCGCGTGCAGCATGAAGTGGCGATGCGCAAC 600
Qy 601 AATGCGATGATCCGCGAGATGCGCAATGTTCCCTTGTGTGAGCTGATGACACTTGG 660
Db 601 AATGCGATGATCCGCGAGATGCGCAATGTTCCCTTGTGTGAGCTGATGACACTTGG 660
Qy 661 AACTGCTTCAGTGGGATCAACACGAGCGCGGCAATTCGCTGACCAACCGCATC 720
Db 661 AACTGCTTCAGTGGGATCAACACGAGCGCGGCAATTCGCTGACCAACCGCATC 720
Qy 721 GTGCAATCGCGGCACTCTTCCCTCAACACTTCCGATGATCTTGGCTACTTACAC 780
Db 721 GTGCAATCGCGGCACTCTTCCCTCAACACTTCCGATGATCTTGGCTACTTACAC 780
Qy 781 GCGCTGAGCGCAGCTGTTGTGCGACATGTGATGACGCGACCTTGCACATCTGGAG 840
Db 781 GCGCTGAGCGCAGCTGTTGTGCGACATGTGATGACGCGACCTTGCACATCTGGAG 840
Qy 841 AAGAACTGCGCGTGCACCGCGCGCTCGAATCTCAAGCCGAGGCGCGCTGCAAG 900
Db 841 AAGAACTGCGCGTGCACCGCGCGCTCGAATCTCAAGCCGAGGCGCGCTGCAAG 900
Qy 901 GACATCGCCATGAGCTCAACGAGATGTACCGGAGTGGACTTGTGAAATACCGCTCC 960
Db 901 GATATCGCATGAACTCAACGAGATGTACCGGAGTGGACTTGTGAAATACCGCTCC 960
Qy 961 TTGGCTATGAGCACTCTTGGCGGTGCTGTGCCACTACTACGTGCGAGGCGCGCTG 1020
Db 961 TTGGCTATGAGCACTCTTGGCGGTGCTGTGCCACTACTACGTGCGAGGCGCGCTG 1020
Qy 1021 GAGCTGCGCGAGGATGACACCGAGCTGAGCCCGGATGCTGTCTCCATGAGCCG 1080
Db 1021 GAACTGCGCGAGGATGACACCGAGCTGAGCCCGGATGCTGTCTCCATGAGCCG 1080
Qy 1081 ATGCTGATGCTGCGCGAGGCAATGCGCGGTGCGCGGCTATCGAGACACATCTTG 1140
Db 1081 ATGCTGATGCTGCGCGAGGCAATGCGCGGTGCGCGGCTATCGAGACACATCTTG 1140
Qy 1141 ATCTGCGGAGAGAGAGCTGCGAGACATCAACCGCTTCCGTTGCTCGGACACAC 1200
Db 1141 ATCTGCGGAGAGAGAGCTGCGAGACATCAACCGCTTCCGTTGCTCGGACACAC 1200
Qy 1201 ATCATCCGCAAC 1212
Db 1201 ATCATCCGCAAC 1212

RESULT 5

US-10-251-078-13

```

; Sequence 13, Application US/10251078
; Publication No. US20030119084A1
; GENERAL INFORMATION:
; APPLICANT: Roche Diagnostics GmbH
; APPLICANT: Shao, Zhixin
; APPLICANT: Schuck, Rainer
; APPLICANT: Kratzsch, Peter
; APPLICANT: Kenkies, Janet
; APPLICANT: Weisser, Harald
; TITLE OF INVENTION: Variants of an Erwinia-type creatinase
; FILE REFERENCE: 20981 EP
; CURRENT APPLICATION NUMBER: US/10/251,078
; CURRENT FILING DATE: 2002-09-20
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 1212
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:variant CT2m10
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1212)
; US-10-251-078-13

```

```

Query Match      80.9%; Score 980; DB 15; Length 1212;
Best Local Similarity 88.0%; Pred. No. 2.5e-257;
Matches 1067; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

```

```

QY 1 ATGACTGACGACATGTTGACGATGATGAAATGACACAGCGCGAGAAATTAATTCGCCG 60
DB 1 ATGACTGACGACATGTTGACGATGATGAAATGACACAGCGCGAGAAATTAATTCGCCG 60
QY 61 TTTTGGATGCGGATGACCCGCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
DB 61 TTTTGGATGCGGATGACCCGCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
QY 121 AATGTCATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
DB 121 AATGTCATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
QY 181 TACTGCTATTTTCGACGCAAGTACGCGATGTCATGACACCAACACGCGACGACGAT 240
DB 181 TACTGCTATTTTCGACGCAAGTACGCGATGTCATGACACCAACACGCGACGACGAT 240
QY 241 TCGGCGCGGCGATGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
DB 241 TCGGCGCGGCGATGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
QY 301 ACCGATGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
DB 301 ACCGATGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
QY 361 CGCATCGGCGATGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
DB 361 CGCATCGGCGATGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
QY 421 CTACCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
DB 421 CTACCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
QY 481 TCGCTCGAAGACGAAAGTGTATTCGCGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
DB 481 TCGCTCGAAGACGAAAGTGTATTCGCGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
QY 541 GCGTCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
DB 541 GCGTCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
QY 601 AATGTCATGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 660
DB 601 AATGTCATGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 660

```

```

QY 661 ACCTGCTTCAAGTGGGCGATCAACACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
DB 661 ACCTGCTTCAAGTGGGCGATCAACACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
QY 721 GTGCAATCCGCGCGACATCTTTTCTGCAACACTTTCGCGATGCTTTCGCGTACAC 780
DB 721 GTGCAATCCGCGCGATCTTTTCTGCAACACTTTCGCGATGCTTTCGCGTACAC 780
QY 781 GCGCTGAGACGACGCTGTTCTGCGACCATGTCATGAGCGCGCGCGCGCGCGCGCGCGCG 840
DB 781 GCGCTGAGACGACGCTGTTCTGCGACCATGTCATGAGCGCGCGCGCGCGCGCGCGCG 840
QY 841 AAGAACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
DB 841 AAGAACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
QY 901 GACATCGCGCGATGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
DB 901 GATATCGCGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
QY 961 TTTGCGCTATGCGCGATCTTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1020
DB 961 TTTGCGCTATGCGCGATCTTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1020
QY 1021 GAGTGGCGGAGGACATGACACCGAGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1080
DB 1021 GAGTGGCGGAGGACATGACACCGAGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1080
QY 1081 ATGTCATGCTGCGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1140
DB 1081 ATGTCATGCTGCGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1140
QY 1141 ATGTCGCGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1200
DB 1141 ATGTCGCGGAGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1200
QY 1201 ATCATCGCGCAAC 1212
DB 1201 ATCATCGCGCAC 1212

```

RESULT 6

US-10-251-078-15

```

; Sequence 15, Application US/10251078
; Publication No. US20030119084A1
; GENERAL INFORMATION:
; APPLICANT: Roche Diagnostics GmbH
; APPLICANT: Shao, Zhixin
; APPLICANT: Schuck, Rainer
; APPLICANT: Kratzsch, Peter
; APPLICANT: Kenkies, Janet
; APPLICANT: Weisser, Harald
; TITLE OF INVENTION: Variants of an Erwinia-type creatinase
; FILE REFERENCE: 20981 EP
; CURRENT APPLICATION NUMBER: US/10/251,078
; CURRENT FILING DATE: 2002-09-20
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 1215
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:variant CT2m28
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1215)
; US-10-251-078-15

```

```

Query Match      80.9%; Score 980; DB 15; Length 1215;
Best Local Similarity 88.0%; Pred. No. 2.5e-257;
Matches 1067; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

```

QY 1 ATGACTGACGACATGTTGCACTGATGAATAATGGCAACAAGCGGAGAAAGATTATGGCCG 60
 Db 1 ATGACTGACGACATGTTGCACTGATGAATAATGGCAACAAGTTGTTTCCCCC 60
 QY 61 TTTTCGGATGCGGAGATGACCGCGCGCAAAAAGAGTTGCGGCTGATGATGCGCAAGAAC 120
 Db 61 TTTTCGGATGCGGAGATGACCGCGCGCAAAAAGAGTTGCGGCTGATGATGCGCAAGAAC 120
 QY 121 AATGTGATGCGGCGCTGTTTACCTCTTATCACTGATCAACTACTATTCGCGCTGCTG 180
 Db 121 GAGGTGACGCTGCGCTGTTTACCTCTTATCACTGATCAACTACTATTCGCTGCTG 180
 QY 181 TACTGCTATTTTGGAGAGCAAGTACGGCATGATGATGACCAACAACCGCAGACAGATT 240
 Db 181 TACTGCTATTTTGGAGAGCAAGTACGGCATGATGATGACCAACAACCGCAGACAGATT 240
 QY 241 TCGGCGGCGGATCGACG 300
 Db 241 TCGGCGGCGGATCGATGCG 300
 QY 301 ACCGACTGCGCGCGCGCGCAATTTCTATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
 Db 301 ACCGACTGCGCGCGCGCGCAATTTCTATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
 QY 361 CGGATGCGGATCGAGTTTCCGACACCTGCACTTCCGCGCGCGCGCGCGCGCGCGCGCG 420
 Db 361 CGGATGCGGATCGAGTTTCCGACACCTGCACTTCCGCGCGCGCGCGCGCGCGCGCGCG 420
 QY 421 CTACCGGCGGCTCGAGTTTCTGCACTACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
 Db 421 CTGCGCGGCGGCTCGAGTTTCTGCACTACGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
 QY 481 TCGCTGGAAGAGCAAGCTGATCGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
 Db 481 TCGCTGGAAGAGCAAGCTGATCGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
 QY 541 GCGTGGCGGCTGCGCATCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
 Db 541 GCGTGGCGGCTGCGCATCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
 QY 601 AATGCGATGATCGCGGAGATGCGCAATCGTTCCCTTCGATGAGTGAACACTGCG 660
 Db 601 AATGCGATGATCGCGGAGATGCGCAATCGTTCCCTTCGATGAGTGAACACTGCG 660
 QY 661 ACTGCTTCAAGTCGGGCAATCAACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
 Db 661 ACTGCTTCAAGTCGGGCAATCAACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
 QY 721 GTGCAATCGGCGCAATCTCTTTCGCTCAACCTTCCGATGATCTTTCGCTCACTAC 780
 Db 721 GTGCAATCGGCGCAATCTCTTTCGCTCAACCTTCCGATGATCTTTCGCTCACTAC 780
 QY 781 GCGCTGAGAGCAAGCTGTTCTGCAACATGATGATGATGATGATGATGATGATGATGAT 840
 Db 781 GCGCTGAGAGCAAGCTGTTCTGCAACATGATGATGATGATGATGATGATGATGATGAT 840
 QY 841 AAGAACTGCGCGCTGATCG 900
 Db 841 AAGAACTGCGCGCTGATCG 900
 QY 901 GACATGCGCATCGAGCTCAACGAGATGATCGCGAGTGGGACTGCTGAGATGATCGCTCC 960
 Db 901 GATATGCGCTTCGAATCAACGAGATGATCGCGAGTGGGACTGCTGAGATGATCGCTCC 960
 QY 961 TTGCGCTATGAGCACTCTTTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
 Db 961 TTGCGCTATGAGCACTCTTTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
 QY 1021 GAGCTGCGGAGGACATCGACACGAGCTGAGACCGCGCGCGCGCGCGCGCGCGCGCGCG 1080
 Db 1021 GAGCTGCGGAGGACATCGACACGAGCTGAGACCGCGCGCGCGCGCGCGCGCGCGCGCG 1080

QY 1081 ATGATGATGCTGCGGAGGCGCATGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1140
 Db 1081 ATGATGATGCTGCGGAGGCGCATGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1140
 QY 1141 ATGCTGCGGAGGAGCGGTGCGCGAGACATCAACCGGCTTCGCTGCGTCCGGAACAAC 1200
 Db 1141 ATGCTGCGGAGGAGCGGTGCGCGAGACATCAACCGGCTTCGCTGCGTCCGGAACAAC 1200
 QY 1201 ATCATCCGCAAC 1212
 Db 1201 ATCATCCGCAAC 1212
 RESULT 7
 US-10-251-078-21
 ; Sequence 21, Application US/10251078
 ; Publication No. US20030119084A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Roche Diagnostics GmbH
 ; APPLICANT: Shao, Zhixin
 ; APPLICANT: Schmuck, Rainer
 ; APPLICANT: Kratzsch, Peter
 ; APPLICANT: Kenkies, Janet
 ; APPLICANT: Weisner, Harald
 ; TITLE OF INVENTION: Variants of an Erwinia-type creatinase
 ; FILE REFERENCE: 20981 EP
 ; CURRENT APPLICATION NUMBER: US/10/251,078
 ; CURRENT FILING DATE: 2002-09-20
 ; NUMBER OF SEQ ID NOS: 32
 ; SOFTWARE: Patentin version 3.1
 ; SEQ ID NO 21
 ; LENGTH: 1212
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: variant Ctsd7
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(1212)
 US-10-251-078-21
 Query Match 80.7%; Score 978.4; DB 15; Length 1212;
 Best Local Similarity 88.0%; Pred. No. 6,7e-257;
 Matches 1066; Conservative 0; Mismatches 146; Indels 0; Gaps 0;
 QY 1 ATGACTGACGACATGTTGCACTGATGAATAATGGCAACAAGCGGAGAAAGATTATGGCCG 60
 Db 1 ATGACTGACGACATGTTGCACTGATGAATAATGGCAACAAGTTGTTTCCCCC 60
 QY 61 TTTTCGGATGCGGAGATGACCGCGCGCAAAAAGAGTTGCGGCTGATGATGCGCAAGAAC 120
 Db 61 TTTTCGGATGCGGAGATGACCGCGCGCAAAAAGAGTTGCGGCTGATGATGCGCAAGAAC 120
 QY 121 AATGTGATGCGGCGCTGTTTACCTCTTATCACTGATCAACTACTATTCGCGCTGCTG 180
 Db 121 GAGGTGACGCTGCGCTGTTTACCTCTTATCACTGATCAACTACTATTCGCTGCTG 180
 QY 181 TACTGCTATTTTGGAGAGCAAGTACGGCATGATGATGACCAACAACCGCAGACAGATT 240
 Db 181 TACTGCTATTTTGGAGAGCAAGTACGGCATGATGATGACCAACAACCGCAGACAGATT 240
 QY 241 TCGGCGGCGGATCGACG 300
 Db 241 TCGGCGGCGGATCGATGCG 300
 QY 301 ACCGACTGCGCGCGCGCGCAATTTCTATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
 Db 301 ACCGACTGCGCGCGCGCGCAATTTCTATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
 QY 361 CGGATGCGGATCGAGTTTCCGACACCTGCACTTCCGCGCGCGCGCGCGCGCGCGCGCG 420
 Db 361 CGGATGCGGATCGAGTTTCCGACACCTGCACTTCCGCGCGCGCGCGCGCGCGCGCGCG 420

```

Oy 421 CTACCGGGGCGTCAAGTTCGTGACATCAAGCAGCCCTCGATGATGAGCGACCATCAAG 480
Db 421 CTGCCCCGGGGTGAAGTTCGTGATCGATGATCGATGATCGATGATCGATGATCGATGAT 480
Oy 481 TGGCTCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Db 481 TGGCTCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Oy 541 GCGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600
Db 541 GCGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600
Oy 601 AATGCGATGATCGGCGAGATCGGCGAGATCGGCGAGATCGGCGAGATCGGCGAGATCGG 660
Db 601 AATGCGATGATCGGCGAGATCGGCGAGATCGGCGAGATCGGCGAGATCGGCGAGATCGG 660
Oy 661 ACCGCGTTCAGTCCGCGCATCAACACCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720
Db 661 ACCGCGTTCAGTCCGCGCATCAACACCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720
Oy 721 GTGCAATCGGCGGAGATCTCTTTCGTGACACCTTTCGTGAGATCTTTCGTGAGATCT 780
Db 721 GTGCAATCGGCGGAGATCTCTTTCGTGACACCTTTCGTGAGATCTTTCGTGAGATCT 780
Oy 781 GCGCTGAGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Db 781 GCGCTGAGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Oy 841 AAGAAAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
Db 841 AAGAAAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
Oy 901 GACATCGGCGGAGAGATCGGCGAGATCGGCGAGATCGGCGAGATCGGCGAGATCGGCG 960
Db 901 GATATCGGCGGAGAGATCGGCGAGATCGGCGAGATCGGCGAGATCGGCGAGATCGGCG 960
Oy 961 TTCGCGTATGAGCGAGATCTCTTTCGTGACACCTTTCGTGAGATCTTTCGTGAGATCT 1020
Db 961 TTCGCGTATGAGCGAGATCTCTTTCGTGACACCTTTCGTGAGATCTTTCGTGAGATCT 1020
Oy 1021 GAGCTGCGGAGAGATCGGCGAGATCGGCGAGATCGGCGAGATCGGCGAGATCGGCGAG 1080
Db 1021 GAGCTGCGGAGAGATCGGCGAGATCGGCGAGATCGGCGAGATCGGCGAGATCGGCGAG 1080
Oy 1081 ATGCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
Db 1081 ATGCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
Oy 1141 ATGCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
Db 1141 ATGCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
Oy 1201 ATCATCGCGCAAC 1212
Db 1201 ATCATCGCGCAAC 1212

```

```

RESULT 8
US-10-251-078-19
; Sequence 19, Application US/10251078
; Publication No. US20030119084A1
; GENERAL INFORMATION:
; APPLICANT: Roche Diagnostics GmbH
; APPLICANT: Shao, Zhixin
; APPLICANT: Schmeck, Rainer
; APPLICANT: Kramsch, Peter
; APPLICANT: Kenkies, Janet
; APPLICANT: Weisner, Harald
; TITLE OF INVENTION: Variants of an Erwinia-type creatinase
; FILE REFERENCE: 20981 EP
; CURRENT APPLICATION NUMBER: US/10/251,078
; CURRENT FILING DATE: 2002-09-20
; NUMBER OF SEQ ID NOS: 32

```

```

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 1212
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: variant CTs62
; NAME/KEY: CDS
; LOCATION: (1)..(1212)
US-10-251-078-19

Query Match      80.6%; Score 976.8; DB 15; Length 1212;
Best Local Similarity 87.9%; Pred. No. 1.8e-256;
Matches 1065; Conservative 0; Mismatches 147; Indels 0; Gaps 0;

Oy 1 ATGACTGACGACATGTTGACGATGAAATGAGCAACAGCGGAGAAAGATTATTCGCGG 60
Db 1 ATGACTGACGACATGTTGACGATGAAATGAGCAACAGCGGAGAAAGATTATTCGCGG 60
Oy 61 TTTTGGATGCGGAGATGACCCCGCCAAAACGAGTTGGCGGCTGAGATGCGCAAGAAC 120
Db 61 TTTTGGATGCGGAGATGACCCCGCCAAAACGAGTTGGCGGCTGAGATGCGCAAGAAC 120
Oy 121 AATGCGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
Db 121 AATGCGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
Oy 181 TACTGCTATTTTCGAGCGAAGTACGAGATGATGATGACGACGACGACGACGAT 240
Db 181 TACTGCTATTTTCGAGCGAAGTACGAGATGATGATGACGACGACGACGAT 240
Oy 241 TGGCGCGGCGATGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
Db 241 TGGCGCGGCGATGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
Oy 301 ACCGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
Db 301 ACCGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
Oy 361 GCGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
Db 361 GCGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
Oy 421 CTACCGGGGCGTGAAGTTCGTGACATCAAGCAGCCCTCGATGATGAGCGACATCAAG 480
Db 421 CTACCGGGGCGTGAAGTTCGTGACATCAAGCAGCCCTCGATGATGAGCGACATCAAG 480
Oy 481 TGGCTCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Db 481 TGGCTCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Oy 541 GCGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600
Db 541 GCGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600
Oy 601 AATGCGATGATCGGCGAGATCGGCGAGATCGGCGAGATCGGCGAGATCGGCGAGATCGG 660
Db 601 AATGCGATGATCGGCGAGATCGGCGAGATCGGCGAGATCGGCGAGATCGGCGAGATCGG 660
Oy 661 ACCGCGTTCAGTCCGCGCATCAACACCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720
Db 661 ACCGCGTTCAGTCCGCGCATCAACACCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720
Oy 721 GTGCAATCGGCGGAGATCTCTTTCGTGACACCTTTCGTGAGATCTTTCGTGAGATCT 780
Db 721 GTGCAATCGGCGGAGATCTCTTTCGTGACACCTTTCGTGAGATCTTTCGTGAGATCT 780
Oy 781 GCGCTGAGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Db 781 GCGCTGAGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Oy 841 AAGAAAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900

```

Query Match	Similarity	6.1%	Score 73.8	DB 19	Length 2733
Best Local	Similarity	44.3%	Pred. No. 6.9e-10		
Matches	398	Conservative	0	Mismatches 452	Indels 9
				Gaps	2
OY	252	CGACGCGCGCCAGCCCTGCGCGCGCAGCTTGCGGACPAACATCACTACACCACTGCGC	311		
Db	2506	CGACCGAGCGCGGTAGCCCGCGGTGGAGCAGCGCTTCCACTTCACTCGGCCCACTTGCC	2447		
OY	312	CCGCGCAAAATTTCTATCGCGCGCGTGCACAGCTGACACAGCGCGCCAAAGCGCATCGCAT	371		
Db	2446	CGAAGACGTGTCTTCTCCGTCAAGGTGCGGTCTCGCTGATCCAAAGCTCATGCGCG	2387		
OY	372	CGAGTTGCAGCAAGCTCAATCTCGACTTCGCGCGCAGCTCGAGGAAGCCCTACCGGGCGT	431		
Db	2386	CGCTACCTCCCCGTGAGGACCTCTCTGTCCGCGAGAGCCCTCGAGCGCAAGCTGCACAT	2327		
OY	432	CGAGTTGCTGCACATAGCCAGCCCTTCGATGTGATGCGCACATCAAGTCGCTGAAAGA	491		

Query Match	5.9%	Score 71.8	DB 20	Length 1262
Best Local Similarity	43.9%	Pred. No. 2.1e-09		
Matches 307, Conservative	0	Mismatches 352	Indels 0	Gaps 0

[illegible]

```

RESULT 11
US-10-437-963-38426
Sequence 38426, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovacic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 38426
LENGTH: 1483
TYPE: DNA
ORGANISM: Oryza sativa

```

```

; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_42062C.1
US-10-437-963-38426

```

Query Match	5.8%	Score 69.8	DB 19	Length 1483
Best Local Similarity	44.9%	Pred. No. 7.6e-09		
Matches 266; Conservative	0	Mismatches 327	Indels 0	Gaps 0

QY	CGATCGGCACACCCAAATGCAATGATGTCCGAGATATGCGCAAAATCGTTCCCTTGTGGAGC	646
Db	309 CGAACCCTGCACGGAGACGGAGATCCGCGCTTCACTGCGCGGAGTACGTCCACTCC	368
QY	647 TGATGGAACCTGGAACCTGTTCCAGTCCGGGCATCAACACGACGGCGGCACAATCCGG	706
Db	369 TCGCGACCTCACCCCGAGAGCTACTTCAACGACCGCGCTTCGGGAGAAAGCGAGG	428
QY	707 TCACCAACCGCATGTGTGCAATCCGGGCAATCCTTTGGCTCAACACTTTCGGATGATCT	766
Db	429 ATGACCAACGGCATTCGGGGGAGACGACATTTGCCCGCTTCGACCGGCTTCGGAAGTACT	488
QY	767 TCGGCTACTTACACCGGCGCTGAGAGGCAACGCTGTTCTGCGACATGTGATGACGCAACGC	826
Db	489 GCGCGGCTACCCCGGCGGCTCTTTCGCGCGGCGCGCGGCTTGTGACGGCGCTTCG	548
QY	827 TCGCATCTTGGAGAAAGAACGTGGCCGTGCATCGCCGCGGACTCGACTGATCAAGCCGG	886
Db	549 ACAATGCCATCAACTGTGTCGGGGGAGATGACACACCGGCTCGGCTGCAAGGCCACCGGCT	608
QY	887 GCGCGCGCTGCAGAGACATGCGCATGTGAGCTCAACGAGATGTACCGGAGTGGGACCTGC	946
Db	609 TCTGCTAACGTCAACGACATGTGCTCGCCATCAACGAGCTCTTCGGCACTTCAAGGGAG	668
QY	947 TGAAGTACCGGCTCCTTGCGCTATGGCGACACTCTTGGCGGTGTGTCACACTACAGCTC	1006
Db	669 TCATCTACGTGCAATGACGCGCACCGGCGAGCGGCTGTGCAAACTCCTTTGTGAAGT	728
QY	1007 GCGAGGCGGCGCTGTGAGCTGCGCGAGGACATGCAACACGAGCTGAAGCCCGGACATGTGTG	1066
Db	729 CCAACGGGGTCAAGAGGTGTCTTTCACCGGTAACGGCAAAATACAGCCGCCACAAAGAACT	788
QY	1067 TCTTCATGAGCCGATGTGTATCTCTGCGGAGGACATGCCGGGTGCCGGCGGCTATGTGCG	1126
Db	789 TCTTCCGGGGCTCCGGCGCCATCAAGAAATCGGCGACGGCGCGGACACTACAGCG	848
QY	1127 AGCAGACATCTCGATGTCTGCGGAGAGACGATCCGAGAACATCACCGGCTTC	1179
Db	849 TGAACGTGCCCCCTTGAACGCGCGGCTCCGCGACACGTTATCAACACGCTGTTC	901

```

RESULT 12
US-10-411-910A-222
; Sequence 222, Application US/10411910A
; Publication No. US20040209256A1
; GENERAL INFORMATION:
; APPLICANT: Dillion, Harrison F.
; TITLE OF INVENTION: Methods and Compositions for Evolving Hydrogenase Genes
; FILE REFERENCE: H2041203-P
; CURRENT APPLICATION NUMBER: US/10/411,910A
; CURRENT FILING DATE: 2003-04-12
; NUMBER OF SEQ ID NOS: 343
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 222
; LENGTH: 1212
; TYPE: DNA
; ORGANISM: Scenedesmus obliquus
; US-10-411-910A-222

```

	Query Match	5.7%;	Score 69.6;	DB 20;	Length 1212;
	Best Local Similarity	42.6%;	Pred. No. 8.3e-09;		
	Matches 363;	Conservative	0;	Mismatches 489;	Indels 0;
Gy	183	CCTGCTATTTGGACGCAGTACGCCATGTTCATCGACCACAACAGCCAGCAGCATTTTC	242		

DB 294 CAGCTGTGCCCCGGCTGGGTGGCCATGATGAGAGAACCCGAGTATCCCTTA 353
QY 243 GGGCGGCATGACGGGGCCAGCCCTGGGGCCGAGCTTGGCGAACAATCACTCAAC 302
DB 354 CCGAGACAGCTGGAGAGCCCGCAGATGATGCTGGGGCCGCTGATCAAGAACTACTTCCG 413
QY 303 CGACTGGGCGCGGCAATTTCTATGCGCGGTGGCCGACGCTGACACAGCGGGCGCAAGCG 362
DB 414 CGCGAGGCGCGGGCCAGCGCCAGGACATCTGCAACGTGAGCGTATGCTGCGTGG 473
QY 363 CATCGGCATGAGTTTGACCACTCAATCTGACTTTCGCGCCGACCTGAGAGAACCTT 422
DB 474 CAAGAGCGGCGAGCGGAGCGCGCGAGCGGACCACTCAACCGCGCGCGCGCGCA 533
QY 423 ACCGGGCGTGAAGTTGTGCAATGAGCCAGCCCTGATGATGATGCGCACTCAAGTC 482
DB 534 CGTGAACAAGTATACCAACCGCCGAGTGGGCAAGATCTTGTGAGAGCGCGCACTCA 593
QY 483 GCTCGAAGAGCAAGACTGATCCGCGAAGGCGCGCGGTGTGATGAGTGGCGCGCGG 542
DB 594 GCTGAAAGAGCTGACAGAGAGCCCTTCAACCCCTGAGGCGAGAGCGAGCGCGCG 653
QY 543 CTGCGGCGCTGCATCAAGCGCGCGCGCGCGAGCATGAAGTGGCGATCGCAACCA 602
DB 654 CCGCTGTTCCGGACCAACCGCGCGCGGTGATGAGGCGCGCGCGCACTGTAAGAGT 713
QY 603 TCGCATGATCCGCGAGATGCGCAATGCTTCCCTTGTGAGAGCTATGACACTGAGC 662
DB 714 GGTGACCGCGAGGCGCTGGGCGCGCGAGCGCAAGCTGACCAACGACCGCTGAGC 773
QY 663 CTGGTTTCAAGTGGGCGATCAACCGAGCGCGCGCAATCCGCTCAACCAACGATCGT 722
DB 774 CCCCCCGCCAGCGCGCCAGCCCGCGCCCGCCAGCGCCCGCCCGCCCGCCCGCC 833
QY 723 GCATTCGCGCAATCTTTCGCTCAACCTTCCGATGATCTTGGCTACTCAACCGC 782
DB 834 CCGTGAAGGCGCAAGACGAGAGCGGATCACTCGAATCGCGCTGAGCGCTGGG 893
QY 783 GCTGAGAGCGCAAGCTGTTTCTGCGACATGTCATGAGCGCAGCTTGAACATCTGGAGAA 842
DB 894 CAACCGCAAGAGCTGATCAAGAGCTGCGCGCGCGAGAGCAAGTACACTTCAACCGA 953
QY 843 GAACGTCGCTGATCGCGCGCGCGCTGAGCTGATCAAGCGCGCGCGCGCTGCAAGGA 902
DB 954 GGTGATGCGCTTCCCGCGCGCTGATCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1013
QY 903 CATCGGCATGAGCTTCAAGAGATGTAACCGAGTGGGACTGCTGAAATACCGCTCTT 962
DB 1014 GATCTGCAAGAGCGCGCAAGCGCGCATGTAACGCTGAGAGCGCGCGCTGATCGCGCG 1073
QY 963 CGGCTATGCGCACTCTTGGGCGGTGCTGCGCACTACGCTGCGGAGCGCGCGTGA 1022
DB 1074 CACCGAAGAACCCCTGATGCGCGCGCTGTAAGAGTTCCTGGGCGAGCGCAACCGCGCA 1133
QY 1023 GCTGCGCGAGGA 1034
DB 1134 CAGGCGCCACGA 1145

RESULT 13
US-10-437-963-57011/c
; Sequence 57011, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Bardazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 57011
; LENGTH: 1169
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_58865C.1
US-10-437-963-57011

Query Match 5 7%; Score 68.6; DB 19; Length 1169;
Best Local Similarity 45.3%; Pred. No. 1.6e-08;
Matches 248; Conservative 0; Mismatches 299; Indels 0; Gaps 0;

QY 512 GCGGCGCGCGCTGCGCGCTGCGCATCAAGCGCGCGCGCGAGCATGAGTGGCATC 591
DB 918 GCGGCG 859
QY 592 GCGACCAACCATGCGATGATCCGAGATGCGCAATGCTTCCCTTGTGAGACTGATG 651
DB 858 AACCTGCG 799
QY 652 GACACCTGAGCCTGTTCCAGTCCGCGCATCAACCGAGCGCGCGCAATCCGCTCAC 711
DB 798 GTGGCGCTCG 739
QY 712 AACCGCATGTCGCAATCCGCGCGCGCGCATCTTTCGCTCAACACTTCCGATGATCTTCCG 771
DB 738 GTGACGT 679
QY 772 TACTAACCGCGCGCTGAGCGCGCGCTGTTCTGCGACATGTGATGAGCGCGCGCTGAGC 831
DB 678 CTGACGCGCGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 619
QY 832 ATCTGGAGAGAAAGTGGCGCGTGTGATCGCGCGCGCGCGCGCGCGCGCGCG 891
DB 618 AGCTTGTGAGAGTGTGATGCGCGCGCTTCTTCAACGAGCTTCAAGCGCGCGCGCG 559
QY 892 CGCTCAAGAGCATCGCATCGAGCTCAACGAGATGTAACCGAGTGGGACTGCTGAG 951
DB 558 ACCACTTCATGCG 499
QY 952 TACCGCTCTTGGCTATGCGCATCTTTCGCGCTGCTGCTGCTGCTGCTGCTGCT 1011
DB 498 CGGCGCTACGCGCGCGCGTGTGATGACTGCTGCTCACTTCCAGTTCTACGCGCTACG 439
QY 1012 GCGCGCGTGAAGCTGCGCGCGAGGACATGCAACCGAGCTGAAGCGCGCGCATGCTGCTCC 1071
DB 438 ACCGAGTGGCGAGCATGATGATGTTTCAAGAGCGAGCGCGCGCAATACCGCGCGCG 379
QY 1072 ATGAGC 1078
DB 378 AAGTGC 372

RESULT 14
US-10-156-761-4932
; Sequence 4932, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761


```

; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 4932
; LENGTH: 834
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURES:
; NAME/KEY: CDS
; LOCATION: (1)..(834)
; US-10-156-761-4932

```

```

Query Match      5.6%; Score 68; DB 15; Length 834;
Best Local Similarity 46.6%; Pred. No. 2.1e-08;
Matches 332; Conservative 0; Mismatches 365; Indels 15; Gaps 3;

```

```

QY 475 ATCAATGCTGCTGGAAGAGCAAGAGCTATCCGCGAAGAGCGCCCGCTGTGTGAGCTGGC 534
DB 10 ATCAAGACCCCGAGAGAGATGCGCAAGATGCTGAGCGCGGCTGTGTGTGCGCCGCTC 69
QY 535 GGGCGGCGCTGCGCGGCTGCTCAAGCGCGGCTGCGCGGATGAGTGGGATGCGC 594
DB 70 CATGCGGCGCACTGATGAGCGCGGCTGCGCGGCGCGCACTGAGAGAGCTGAGAGAGTC 129
QY 595 ACCACCAATGCAATGATCCGCGAGATGCGCAATTCCTTCCCTTGTGAGCTGATGAC 654
DB 130 GCGCGGAGAGGTGCTCGCGAGAGCAAGGTGCGCAAGTCTGCTCGGCTGCGCGGCTTC 189
QY 655 ACCTGAGCTGCTGCTGAGTGGGATCAACACGAGCGGCGGCAATCCGCTGACCAAC 714
DB 190 CCGGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 249
QY 715 CCGATGCTGCAATCCGCGGCACTCTTGTGCTCAACCTTCCGATGATCTTGTGCTG 774
DB 250 ACCGCTCTGAAAGAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 309
QY 775 TACACGCGCTGAGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 811
DB 310 CAGCGGAGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 369
QY 832 ATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 891
DB 370 CTCTCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 429
QY 892 CGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 951
DB 430 CGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 489
QY 952 TACCGCTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1011
DB 490 GCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 543
QY 1012 GCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1071
DB 544 ATGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 603
QY 1072 ATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1131
DB 604 CTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 663
QY 1132 GACATCTGATGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1183
DB 664 GA-----GCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 709

```

```

RESULT 15
US-10-363-345A-31407/c
; Sequence 31407, Application US/1033345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; FILE OF INVENTION: cytosines in genomic DNA in the sequence context of 5'-CpG-3
; FILE REFERENCE: 801/1227
; CURRENT APPLICATION NUMBER: US/10/363,345A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 31407
; LENGTH: 985
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURES:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-Island No: 31407
; US-10-363-345A-31407

```

```

Query Match      5.6%; Score 68; DB 20; Length 985;
Best Local Similarity 43.0%; Pred. No. 2.2e-08;
Matches 385; Conservative 0; Mismatches 510; Indels 1; Gaps 1;

```

```

QY 152 ACTGATCACTAATTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 211
DB 937 ACATGACATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 878
QY 212 TATGACCAACCAAGCGGCAAGATTTGCGCGGATTCGAGCGGCGGCGGCGGCGGCG 271
DB 877 ACATCTCACTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 818
QY 272 GCGGAGCTTGGCGGCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 331
DB 817 ACATGACATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 758
QY 332 CCGTGGC-CAGCTGCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 390
DB 757 ACATCTCACTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 698
QY 391 CTGATCTTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 450
DB 697 TACATGACATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 638
QY 451 CAGCCCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 510
DB 637 CCGCTTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 578
QY 511 GCGGCGGCGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 570
DB 577 AACGTCAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 518
QY 571 CCGGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 630
DB 517 GAGCTTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 458
QY 631 TTCCCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 690
DB 457 TACATCTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 398
QY 691 GCGGCGCAATCCGCTGCAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 750
DB 397 AACGTCAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 338
QY 751 ACCTTCCGAGATCTTGGCTTACTCAACCGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 810
DB 811 GTCGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 870
QY 277 TACGTCAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 218
DB 217 GACATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 158

```


Qy 931 CCGAGTGGGACCTGCTGAAGTACCGCTCCTTGGCTATGCGCACCTCCCTTGGCGGCTG 990
Db 157 TACGTGACGCTCTACATCTACGTGACGTGACGTCTACATCAACATCAACGTGACGTC 98
Qy 991 TGGCACTACTACGCTGCGAGGCCGCGCTGAGACTGCGAGGACATCGACCGA 1046
Db 97 TACATCAACATCTACATCAACGTGACGTGACGTCTACATCAACGTGACGCTTA 42

Search completed: July 8, 2005, 01:32:51
Job time : 839 secs

Query Match	100.0%;	Score 2190;	DB 2;	Length 404;
Best Local Similarity	100.0%;	Pred. No. 1.5e-214;		

```
Matches 404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MTDDMLHVMKMHNGEKDYSFSDAEMTRONDVGMMAKNNDALFTSYHCINYSGL 60
DB 1 MTDDMLHVMKMHNGEKDYSFSDAEMTRONDVGMMAKNNDALFTSYHCINYSGL 60
QY 61 YCYFGRKYGWVIDHNNATTSISAGIDGQPMRRSFGDNITYTDWRDNFRAVQLTGAK 120
DB 61 YCYFGRKYGWVIDHNNATTSISAGIDGQPMRRSFGDNITYTDWRDNFRAVQLTGAK 120
QY 121 RIGIEFDHNLDFRRQLEALPGVEFVDISQPSMMNRITKSLEQKIREGARVCVGA 180
DB 121 RIGIEFDHNLDFRRQLEALPGVEFVDISQPSMMNRITKSLEQKIREGARVCVGA 180
QY 181 ACAAAIKAGVBEHEVAIATTNAMIREIAKSPFVELMDTWTFQSGINTDGANPTNRI 240
DB 181 ACAAAIKAGVBEHEVAIATTNAMIREIAKSPFVELMDTWTFQSGINTDGANPTNRI 240
QY 241 VQSGDILSLNTPFMITFGYTTALERTLFCDHVDDASLDIWEKNVAVHRRGLELIPGARCK 300
DB 241 VQSGDILSLNTPFMITFGYTTALERTLFCDHVDDASLDIWEKNVAVHRRGLELIPGARCK 300
QY 301 DIAIELNEMVREMDLKYRSFGYHSGFVLCYHGREAGVELREDIDTELKPGVVSMEP 360
DB 301 DIAIELNEMVREMDLKYRSFGYHSGFVLCYHGREAGVELREDIDTELKPGVVSMEP 360
QY 361 MVMLEPGMPGAGGYREHDILIVGEDGAEINTGPPFGEHNIIRN 404
DB 361 MVMLEPGMPGAGGYREHDILIVGEDGAEINTGPPFGEHNIIRN 404

RESULT 2
ID AAM22893 standard; protein; 404 AA.
XX
AC AAM22893;
XX
DT 17-OCT-2003 (revised)
DT 02-MAR-1998 (first entry)
XX
DE Createine amidinohydroxylase enzyme.
XX
KW Createine amidinohydroxylase enzyme; sarcosine; urea; dye; Km.
XX
OS Alcaligenes faecalis; - strain TE3581 (FERM P-14237).
XX
FH Key Location/Qualifiers
FT Misc-difference 145 /note= "encoded by GAC"
XX
XX BEP90303-A1.
XX
XX 20-AUG-1997.
XX
XX PF 13-FEB-1997; 97EP-00102270.
XX
XX PR 13-FEB-1996; 96JP-00025435.
XX
XX PA (TOYM ) TOYO BOSEKI KK.
XX
XX PI Sogabe A, Hattori T, Nishiya Y, Kawamura Y;
XX
XX WPI; 1997-404731/38.
XX
XX DR N-PSDB; AAT38807.
XX
XX PT Createine amidinohydroxylase enzyme with low Km - for use in assay for
XX creatine.
XX
XX Disclosure; Page 13-14; 21pp; English.
XX
XX A novel creatine amidinohydroxylase enzyme has been developed which
XX catalyses the reaction of creatine with water to form sarcosine and urea,
XX at stable at temperatures of up to 50 degrees Celsius (pH 7.5, 30
```

```
CC minutes) and has an optimum temperature of 40-50 degrees Celsius, optimum
CC pH of 8-9, a Km value for creatine of 3.5-10 mM in a binding assay using
CC sarcosine oxidase and peroxidase, a molecular weight of 43 kD by SDS-PAGE
CC and an isoelectric point of 3.5. The present sequence represents creatine
CC amidinohydroxylase derived from Alcaligenes faecalis strain TE3581 (FERM P-
CC 14237), which is the wild type creatine amidinohydroxylase to be mutated in
CC the present invention. The enzyme can be used to determine creatine in a
CC sample by measuring the absorbance of a dye formed by reacting the sample
CC with a reagent, comprising the enzyme, sarcosine oxidase and a
CC composition for detecting hydrogen peroxide, e.g. for diagnosis of
CC uraemia, chronic nephritis, gigantism and tonic muscular dystrophy. The
CC enzyme has a lower Km value than prior art creatine amidinohydroxylase (cf.
CC US 5451520). (Updated on 17-OCT-2003 to standardise OS field)
XX
SQ Sequence 404 AA;
Query Match 100.0%; Score 2190; DB 2; Length 404;
Best Local Similarity 100.0%; Pred. No. 1.5e-214;
Matches 404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MTDDMLHVMKMHNGEKDYSFSDAEMTRONDVGMMAKNNDALFTSYHCINYSGL 60
DB 1 MTDDMLHVMKMHNGEKDYSFSDAEMTRONDVGMMAKNNDALFTSYHCINYSGL 60
QY 61 YCYFGRKYGWVIDHNNATTSISAGIDGQPMRRSFGDNITYTDWRDNFRAVQLTGAK 120
DB 61 YCYFGRKYGWVIDHNNATTSISAGIDGQPMRRSFGDNITYTDWRDNFRAVQLTGAK 120
QY 121 RIGIEFDHNLDFRRQLEALPGVEFVDISQPSMMNRITKSLEQKIREGARVCVGA 180
DB 121 RIGIEFDHNLDFRRQLEALPGVEFVDISQPSMMNRITKSLEQKIREGARVCVGA 180
QY 181 ACAAAIKAGVBEHEVAIATTNAMIREIAKSPFVELMDTWTFQSGINTDGANPTNRI 240
DB 181 ACAAAIKAGVBEHEVAIATTNAMIREIAKSPFVELMDTWTFQSGINTDGANPTNRI 240
QY 241 VQSGDILSLNTPFMITFGYTTALERTLFCDHVDDASLDIWEKNVAVHRRGLELIPGARCK 300
DB 241 VQSGDILSLNTPFMITFGYTTALERTLFCDHVDDASLDIWEKNVAVHRRGLELIPGARCK 300
QY 301 DIAIELNEMVREMDLKYRSFGYHSGFVLCYHGREAGVELREDIDTELKPGVVSMEP 360
DB 301 DIAIELNEMVREMDLKYRSFGYHSGFVLCYHGREAGVELREDIDTELKPGVVSMEP 360
QY 361 MVMLEPGMPGAGGYREHDILIVGEDGAEINTGPPFGEHNIIRN 404
DB 361 MVMLEPGMPGAGGYREHDILIVGEDGAEINTGPPFGEHNIIRN 404

RESULT 3
ID AAM61905 standard; protein; 404 AA.
XX
XX AAM61905;
XX
XX AC AAM61905;
XX
XX DT 16-SEP-1998. (first entry)
XX
XX DE Stable creatine amidinohydroxylase enzyme.
XX
XX KW Createine amidinohydroxylase; mutant; stable; enzyme; diagnostic agent.
XX
XX OS Alcaligenes faecalis.
XX
XX FH Key Location/Qualifiers
XX FT Misc-difference 145 /note= "encoded by GAC"
XX
XX PN JP10174585-A.
XX
XX PD 30-JUN-1998.
XX
XX PF 17-DEC-1996; 96JP-00337027.
XX
```

PR 17-DEC-1996; 96JP-00337027.

XX (TOYM) TOYOBO KK.

XX WPI; 1998-421167/36.

DR N-PSDB; AAV35699.

PT New creatine amidino-hydrazase used as diagnostic agent - is more stable in
PT neutral buffer than wild type creatine amidino-hydrazase.

PS Claim 5; Page 11-13, 14pp; Japanese.

XX This represents a stable creatine amidinohydrazase which is a mutant
CC creatine amidinohydrazase and has improved long-term stability in a neutral
CC buffer compared to wild type creatine amidinohydrazase. A recombinant
CC plasmid containing the stable creatine amidinohydrazase gene can be used to
CC transform a cell for the recombinant production of the enzyme. This
CC stable creatine amidinohydrazase is useful as a diagnostic agent can be
CC produced commercially

XX Sequence 404 AA;

Query Match 100.0%; Score 2190; DB 2; Length 404;
Best Local Similarity 100.0%; Pred. No. 1.5e-214;
Matches 404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MTDDMLHWKMHNGEKDYSPSPDAEMTRONDVRGMANKNDALFTSYHCINYYSGWL 60
DB 1 MTDDMLHWKMHNGEKDYSPSPDAEMTRONDVRGMANKNDALFTSYHCINYYSGWL 60
QY 61 YCFGRKGYMVIDHNNATTISAGIDGQPMWRSPGDNITYTDMRDNFYRAVROLTTGAK 120
DB 61 YCFGRKGYMVIDHNNATTISAGIDGQPMWRSPGDNITYTDMRDNFYRAVROLTTGAK 120
QY 121 RIGIEFDHVNLDPRQLEALPGVEFVDISQPSMMWRITKSLEOKLIREGARVCDVGGA 180
DB 121 RIGIEFDHVNLDPRQLEALPGVEFVDISQPSMMWRITKSLEOKLIREGARVCDVGGA 180
QY 181 ACAAIRKAGVPEHEVAIATTNAMIREIAKSPPEVLEMDTWTFOSGINTDGAHNVTVRI 240
DB 181 ACAAIRKAGVPEHEVAIATTNAMIREIAKSPPEVLEMDTWTFOSGINTDGAHNVTVRI 240
QY 241 VOSGDILSLNTPPMIFGYTTALERTLFCDHVDASLDIWEKNVAHVRGSELIRKGARCK 300
DB 241 VOSGDILSLNTPPMIFGYTTALERTLFCDHVDASLDIWEKNVAHVRGSELIRKGARCK 300
QY 301 DIAELNEMRYEMWDLKTYRSFGYGHSGFVLCHYYGREAGVELREDIDTELKRGVVSMEP 360
DB 301 DIAELNEMRYEMWDLKTYRSFGYGHSGFVLCHYYGREAGVELREDIDTELKRGVVSMEP 360
QY 361 MVMLEPGMGAGGYREHDILIVGEDGAENITGPPGPEHNIIRN 404
DB 361 MVMLEPGMGAGGYREHDILIVGEDGAENITGPPGPEHNIIRN 404

```

RESULT 4

AAB09976 standard; protein; 404 AA.

XX AAB09976;

DT 19-OCT-2000 (first entry)

XX AlcaIlgenees sp. creatine amidinohydrolase protein.

XX Creatine amidinohydrolase; thermostable; diagnosis; kidney disease;

KM bllirubin.

OS AlcaIlgenees sp.

PN W0200031245-A1.

PD 02-JUN-2000.

XX 25-NOV-1999; 99MO-JP006583.

XX 25-NOV-1998; 98JP-00334252.

XX (KIKK) KIRKOMAN CORP.

XX Furukawa K, Ichikawa T;

XX WPI; 2000-411951/35.

XX Highly thermostable creatine amidinohydrolase with optimum pH in weakly
PT acidic region, useful in assaying serum or urine creatine for diagnosis
PT of e.g. kidney diseases, scarcely affected by bilirubin.

PS Claim 3; Page 19-20, 23pp; Japanese.

XX This invention describes a novel highly thermostable creatine
CC amidinohydrolase (I) with optimum pH in a weakly acidic region, useful in
CC assaying serum or urine creatine for diagnosis of e.g. kidney diseases,
CC and scarcely affected by bilirubin. The enzyme is produced by a
CC transformant Escherichia coli in high efficiency. Due to the enzyme
CC having an optimum pH value in the weakly acidic region, it is scarcely
XX affected by bilirubin when assaying creatine

XX Sequence 404 AA;

Query Match 100.0%; Score 2190; DB 3; Length 404;
Best Local Similarity 100.0%; Pred. No. 1.5e-214;
Matches 404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MTDDMLHWKMHNGEKDYSPSPDAEMTRONDVRGMANKNDALFTSYHCINYYSGWL 60
DB 1 MTDDMLHWKMHNGEKDYSPSPDAEMTRONDVRGMANKNDALFTSYHCINYYSGWL 60
QY 61 YCFGRKGYMVIDHNNATTISAGIDGQPMWRSPGDNITYTDMRDNFYRAVROLTTGAK 120
DB 61 YCFGRKGYMVIDHNNATTISAGIDGQPMWRSPGDNITYTDMRDNFYRAVROLTTGAK 120
QY 121 RIGIEFDHVNLDPRQLEALPGVEFVDISQPSMMWRITKSLEOKLIREGARVCDVGGA 180
DB 121 RIGIEFDHVNLDPRQLEALPGVEFVDISQPSMMWRITKSLEOKLIREGARVCDVGGA 180
QY 181 ACAAIRKAGVPEHEVAIATTNAMIREIAKSPPEVLEMDTWTFOSGINTDGAHNVTVRI 240
DB 181 ACAAIRKAGVPEHEVAIATTNAMIREIAKSPPEVLEMDTWTFOSGINTDGAHNVTVRI 240
QY 241 VOSGDILSLNTPPMIFGYTTALERTLFCDHVDASLDIWEKNVAHVRGSELIRKGARCK 300
DB 241 VOSGDILSLNTPPMIFGYTTALERTLFCDHVDASLDIWEKNVAHVRGSELIRKGARCK 300
QY 301 DIAELNEMRYEMWDLKTYRSFGYGHSGFVLCHYYGREAGVELREDIDTELKRGVVSMEP 360
DB 301 DIAELNEMRYEMWDLKTYRSFGYGHSGFVLCHYYGREAGVELREDIDTELKRGVVSMEP 360
QY 361 MVMLEPGMGAGGYREHDILIVGEDGAENITGPPGPEHNIIRN 404
DB 361 MVMLEPGMGAGGYREHDILIVGEDGAENITGPPGPEHNIIRN 404

```

RESULT 5

AAU08727 standard; protein; 404 AA.

XX AAU08727;

DT 19-DEC-2001 (first entry)

XX Creatine amidinohydrolase polypeptide.

XX Creatine amidinohydrolase; water; sarcosine; urea; creatinine; uraemia;

KM chronic nephritis; acute nephritis; tonic muscular dystrophy; gigantism;

KM pigment absorbance.

XX OS Alcaligenes faecalis.
XX PN BP132467-A2.
XX PD 12-SEP-2001.
XX PF 13-FEB-1997; 2001EP-00113052.
XX PR 13-FEB-1996; 96JP-00025435.
XX PR 13-FEB-1997; 97EP-00102270.
XX PA (TOYM) TOYO BOSEKI KK.
XX PI Sogabe A, Hattori T, Nishiya Y, Kawamura Y;
XX DR WPI; 2001-612481/71.
XX DR N-PSDB; AAS14742.
XX PT New creatine amidinohydrolase, useful as a routine reagent for clinical
XX PT tests for determining creatine and creatinine in biological samples,
XX PT particularly useful in diagnosing diseases such as uremia or chronic
XX PT nephritis.
XX PS Disclosure; Page 14-15; 21pp; English.
XX CC The invention relates to Alcaligenes faecalis creatine amidinohydrolase,
XX CC which catalyses creatine and water to sarcosine and urea. Creatine
XX CC amidinohydrolase can be produced by culturing a microorganism producing
XX CC the protein in a nutrient medium and recovering the protein from the
XX CC resulting culture. Creatine amidinohydrolase is useful as a routine
XX CC reagent for clinical tests for determining creatine and creatinine in
XX CC biological samples. This is particularly useful in diagnosing diseases
XX CC such as uremia, chronic nephritis, acute nephritis, glaucoma and tonic
XX CC muscular dystrophy. The presence of creatine in a sample can be
XX CC determined by measuring an absorbance of a pigment produced by the
XX CC reaction of a reagent containing creatine amidinohydrolase with the
XX CC sample. This sequence represents Alcaligenes faecalis creatine
XX CC amidinohydrolase polypeptide
XX CC
XX SQ Sequence 404 AA;

Query Match 100.0%; Score 2190; DB 4; Length 404;
Best Local Similarity 100.0%; Pred. No. 1.5e-214;
Matches 404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTDDMLHWKMHNGEKDYSFSDAEMTRRQNDVRGMANKNVDAALFTSYHCINYSGL 60
DB 1 MTDDMLHWKMHNGEKDYSFSDAEMTRRQNDVRGMANKNVDAALFTSYHCINYSGL 60
QY 61 YCYFGRKYGWVIDHNNATTISAGIDGQPMWRSPFGDNITYTDMWRDNFRAVRQLTTGAK 120
DB 61 YCYFGRKYGWVIDHNNATTISAGIDGQPMWRSPFGDNITYTDMWRDNFRAVRQLTTGAK 120
QY 121 RIGIEFDHVNLDPRROLEALPGVEFVDISQPSMMWRRTIKSLBEQKLREGARVCDVGGA 180
DB 121 RIGIEFDHVNLDPRROLEALPGVEFVDISQPSMMWRRTIKSLBEQKLREGARVCDVGGA 180
QY 181 ACAAAIKAGVPEHEVAIATTNAMIIRIAKSPFVELMDTMTWFOGINTDGAHNPVTNRI 240
DB 181 ACAAAIKAGVPEHEVAIATTNAMIIRIAKSPFVELMDTMTWFOGINTDGAHNPVTNRI 240
QY 241 VOSGDILSLNTPPMIFGYTTALERTLFCDHVDASLDIWEKNVAVHRRGLELIKPGARCK 300
DB 241 VOSGDILSLNTPPMIFGYTTALERTLFCDHVDASLDIWEKNVAVHRRGLELIKPGARCK 300
QY 301 DIAIENEMYREMDLTKRSPGYGHSFGVLCYHYGREGAVELREDIDTELKPGMVSMER 360
DB 301 DIAIENEMYREMDLTKRSPGYGHSFGVLCYHYGREGAVELREDIDTELKPGMVSMER 360
QY 361 MVMLEPGMGAGYREHDLIVGEDGAENITGFPFPEHNITRN 404
DB 361 MVMLEPGMGAGYREHDLIVGEDGAENITGFPFPEHNITRN 404

RESULT 6
AAM51471
ID AAM51471 standard; protein; 404 AA.
XX AC AAM51471;
XX AC 07-AUG-2003 (revised)
XX DT 28-JAN-2002 (first entry)
XX DE Alcaligenes faecalis creatineamidinohydrolase.
XX KW Alcaligenes faecalis; TE3581; FERM P14237; creatineamidinohydrolase.
XX OS Alcaligenes faecalis.
XX FH Key Location/Qualifiers
XX FH Misc-difference 145 /note= "Encoded by GAC"
XX FT JP2001252088-A.
XX PN 18-SEP-2001.
XX PD 16-MAY-1995; 2001JP-00051054.
XX PF 16-MAY-1995; 95JP-00117283.
XX PR 16-MAY-1995; 95JP-00117283.
XX PA (TOYM) TOYOBO KK.
XX DR WPI; 2002-003140/01.
XX DR N-PSDB; AA199856.
XX PS A gene encoding creatineamidinohydrolase.
XX PT Claim 1; Page 9-10; 11pp; Japanese.
XX CC The invention relates to Alcaligenes faecalis TE3581 (FERM P-14237)
XX CC creatineamidinohydrolase and the encoding gene. The gene can be used for
XX CC the commercial preparation of creatineamidinohydrolase. (Updated on 07-
XX CC AUG-2003 to correct OS field.)
XX CC
XX SQ Sequence 404 AA;

Query Match 100.0%; Score 2190; DB 5; Length 404;
Best Local Similarity 100.0%; Pred. No. 1.5e-214;
Matches 404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTDDMLHWKMHNGEKDYSFSDAEMTRRQNDVRGMANKNVDAALFTSYHCINYSGL 60
DB 1 MTDDMLHWKMHNGEKDYSFSDAEMTRRQNDVRGMANKNVDAALFTSYHCINYSGL 60
QY 61 YCYFGRKYGWVIDHNNATTISAGIDGQPMWRSPFGDNITYTDMWRDNFRAVRQLTTGAK 120
DB 61 YCYFGRKYGWVIDHNNATTISAGIDGQPMWRSPFGDNITYTDMWRDNFRAVRQLTTGAK 120
QY 121 RIGIEFDHVNLDPRROLEALPGVEFVDISQPSMMWRRTIKSLBEQKLREGARVCDVGGA 180
DB 121 RIGIEFDHVNLDPRROLEALPGVEFVDISQPSMMWRRTIKSLBEQKLREGARVCDVGGA 180
QY 181 ACAAAIKAGVPEHEVAIATTNAMIIRIAKSPFVELMDTMTWFOGINTDGAHNPVTNRI 240
DB 181 ACAAAIKAGVPEHEVAIATTNAMIIRIAKSPFVELMDTMTWFOGINTDGAHNPVTNRI 240
QY 241 VOSGDILSLNTPPMIFGYTTALERTLFCDHVDASLDIWEKNVAVHRRGLELIKPGARCK 300
DB 241 VOSGDILSLNTPPMIFGYTTALERTLFCDHVDASLDIWEKNVAVHRRGLELIKPGARCK 300
QY 301 DIAIENEMYREMDLTKRSPGYGHSFGVLCYHYGREGAVELREDIDTELKPGMVSMER 360
DB 301 DIAIENEMYREMDLTKRSPGYGHSFGVLCYHYGREGAVELREDIDTELKPGMVSMER 360

OY 361 MWLPEGMPGAGYREHDLIVGEDGAEINTGPPGPEHNIIRN 404
 DB 361 MWLPEGMPGAGYREHDLIVGEDGAEINTGPPGPEHNIIRN 404

RESULT 7

AB05660
 ID AB05660 standard; protein; 404 AA.
 XX
 AC AB05660;

XX 29-APR-2002 (first entry)

DE Creatine amidinohydrazase protein SEQ ID NO:1.

XX Creatine amidinohydrazase; enzyme; mutant; stable; clinical diagnosis.

XX Alcalligenes faecalis.

OS
 FH Key Location/Qualifiers

FT MISC-difference 145 /note="encoded by GAC"

XX JP2001346594-A.

XX 18-DEC-2001.

XX 17-DEC-1996; 2001JP-00121708.

XX 17-DEC-1996; 96JP-00337027.

XX (TOYM) TOYORO KK.

XX WPI: 2002-145187/19.

XX N-PSDB; ABA93696.

DR Mutant creatine amidinohydrazase for use in clinical diagnosis has long-term stability in a neutral buffer solution.

PS Claim 5; Page 10-11; 15pp; Japanese.

XX The present invention describes a stable mutant creatine amidinohydrazase having a long-term stability in a neutral buffer solution. Also described are: (1) a creatine amidinohydrazase gene encoding the above stable creatine amidinohydrazase; (2) a gene encoding a mutant creatine amidinohydrazase having a long-term stability in a neutral buffer solution compared to wild type creatine amidinohydrazase, in which at least one gene among those found in a fully defined 1212 nucleotide sequence (see CC ABA93696) is replaced by another gene; (3) a recombinant plasmid containing a gene encoding the above creatine amidinohydrazase; (4) a cell transformed by the above plasmid; and (5) a method for the preparation of a stable creatine amidinohydrazase in which the above cell is cultured in a medium and creatine amidinohydrazase is collected. The creatine amidinohydrazase is useful as a clinical diagnosing agent. The present CC sequence represents a creatine amidinohydrazase protein sequence given in CC the present invention

XX Sequence 404 AA;

Query Match 100.0%; Score 2190; DB 5; Length 404;
 Best Local Similarity 100.0%; Pred. No. 1.5e-214;
 Matches 404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MTDDMLHYMKHNGEKDYPSPSDAEMTRONDVRGMAKNVDAALFTSYHCINITYSGWL 60
 DB 1 MTDDMLHYMKHNGEKDYPSPSDAEMTRONDVRGMAKNVDAALFTSYHCINITYSGWL 60
 OY 61 YCYFGRKKGWVTDHNNATTISAGIDGQPMWRSPGDNITTYDWRDNYFRAVROLTTGAK 120
 DB 61 YCYFGRKKGWVTDHNNATTISAGIDGQPMWRSPGDNITTYDWRDNYFRAVROLTTGAK 120
 OY 121 RIGIFDHVNLDFRRQLSEALPGVEFVDISQSPMMWRITKSLSEQKLIREGARVCDVGGA 180
 DB 121 RIGIFDHVNLDFRRQLSEALPGVEFVDISQSPMMWRITKSLSEQKLIREGARVCDVGGA 180

DB 121 RIGIFDHVNLDFRRQLSEALPGVEFVDISQSPMMWRITKSLSEQKLIREGARVCDVGGA 180
 OY 181 ACAAAIKAGVPEHEVAIATTNMIKEIASPPVEIMDTWTFQSGINTDGANPYTNRI 240
 DB 181 ACAAAIKAGVPEHEVAIATTNMIKEIASPPVEIMDTWTFQSGINTDGANPYTNRI 240
 OY 241 VOSGDILSINTFPMIFGYTTALERTLFCDHVDASIDIEKXNAVHRRGLEIKPGARCK 300
 DB 241 VOSGDILSINTFPMIFGYTTALERTLFCDHVDASIDIEKXNAVHRRGLEIKPGARCK 300
 OY 301 DIAIEINEMRYREMDLKTFRSFGYHSGFVLCYHGRACVELEIDITELKGMVVSMBP 360
 DB 301 DIAIEINEMRYREMDLKTFRSFGYHSGFVLCYHGRACVELEIDITELKGMVVSMBP 360
 OY 361 MWLPEGMPGAGYREHDLIVGEDGAEINTGPPGPEHNIIRN 404
 DB 361 MWLPEGMPGAGYREHDLIVGEDGAEINTGPPGPEHNIIRN 404

RESULT 8

AB05660
 ID AAR94463 standard; protein; 404 AA.

XX AAR94463;

XX 16-OCT-2003 (revised)

XX 26-SEP-1996 (first entry)

DE Creatine amidinohydrolase.

XX Creatine amidinohydrolase; CAH; kidney; disease; Alcalligenes.

OS Alcalligenes sp; KS-85 FERM BP-4487.

XX DE19536506-AI.

XX 04-APR-1996.

XX 29-SEP-1995; 95DE-01036506.

XX 29-SEP-1994; 94JP-00235737.

XX (KIRK) KIRKMAN CORP.

PI Furukawa K, Ichikawa T, Suzuki M, Koyama Y;

DR WPI: 1996-180805/19.

DR N-PSDB; AAT13291.

PT DNA encoding creatine amidinohydrolase - useful for quantification of PT creatine to, e.g. diagnose kidney disease.

PS Claim 1; Page 11-13; 18pp; German.

XX CAH is used for quantification of creatine, e.g. to diagnose kidney CC disease by measuring creatine content of serum or urine. CAH DNA can be CC inserted into host cells for the prodn. of CAH. CAH can now be produced CC efficiently without having to add creatine to the culture medium.
 CC (Updated on 16-OCT-2003 to standardise OS field)

XX Sequence 404 AA;

Query Match 98.1%; Score 2149; DB 2; Length 404;
 Best Local Similarity 98.0%; Pred. No. 2.3e-210;
 Matches 396; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 1 MTDDMLHYMKHNGEKDYPSPSDAEMTRONDVRGMAKNVDAALFTSYHCINITYSGWL 60
 DB 1 MTDDMLHYMKHNGEKDYPSPSDAEMTRONDVRGMAKNVDAALFTSYHCINITYSGWL 60
 OY 61 YCYFGRKKGWVTDHNNATTISAGIDGQPMWRSPGDNITTYDWRDNYFRAVROLTTGAK 120
 DB 61 YCYFGRKKGWVTDHNNATTISAGIDGQPMWRSPGDNITTYDWRDNYFRAVROLTTGAK 120

QY 121 RIGIEFDHVNLDPRQLEENLPGVEFVDISQPSMMRTIKSLBEOKLIREGARVCVGA 180
 DB 122 RIGIEFDHVNLDPRQLEENLPGVEFVDISQPSMMRTIKSLBEOKLIREGARVCVGA 180
 QY 181 ACAAAIKAGVPEHEVAIATTNAMIIRIAKSPFVELMDTWTWFOGINTDGAHPVTNRI 240
 DB 181 ACAAAIKAGVPEHEVAIATTNAMIIRIAKSPFVELMDTWTWFOGINTDGAHPVTNRI 240
 QY 241 VOSGDILSINTFPMIFGYTALERTLFCDHVDDASLDIWEKNVAVHRRGLELIKPGARCK 300
 DB 241 VOSGDILSINTFPMIFGYTALERTLFCDHVDDASLDIWEKNVAVHRRGLELIKPGARCK 300
 QY 301 DIAIELNEMRYEMDILKTRSPGYGHSFGVLCYHGREAGVELREDIDTELKPGMVSMMP 360
 DB 301 DIAIELNEMRYEMDILKTRSPGYGHSFGVLCYHGREAGVELREDIDTELKPGMVSMMP 360
 QY 361 MWMLPEGMPGAGGYREHDLIVGEDGAENITGPPGPEHNIIRN 404
 DB 361 MWMLPEGMPGAGGYREHDLIVGEDGAENITGPPGPEHNIIRN 404

RESULT 9

AAB12775
 ID AAB12775 standard; protein; 404 AA.

AC AAB12775;

DT 22-NOV-2000 (first entry)

DE Alcaligenes thermostable creatine amidinohydrolase protein SEQ ID NO.1.

KW Alcaligenes; thermostable creatine amidinohydrolase; kidney disease.

OS Alcaligenes sp.

PN WO200040708-A1.

PD 13-JUL-2000.

PF 28-DEC-1999; 99MO-JP007424.

PR 01-JAN-1999; 99JP-00033359.

PA (KIKK) KIKKOMAN CORP.

PI Furukawa K, Koyama Y, Suzuki M;

DR WPI; 2000-475827/41.

PT Novel thermostable Alcaligenes-derived creatine amidinohydrolase, useful for the diagnosis of kidney diseases and related diseases.

PS Claim 4; Page 20-22; 24pp; Japanese.

CC The present sequence represents a thermostable creatine amidinohydrolase isolated from Alcaligenes sp.. The thermostable creatine amidinohydrolase (I) that hydrolyses 1 mol of creatine to give 1 mol of urea, has a substrate specificity to creatine, has an optimum pH range of 7-8, has a stable pH range of 4-11, has an optimum operating temperature of 45 plus degrees Celsius, is stable at 53 plus degrees Celsius, and has a molecular weight of 9200 Da as determined by the gel filtration method. CC The enzyme is applicable in diagnosis of kidney diseases and related diseases

SO Sequence 404 AA;

Query Match 98.1%; Score 2149; DB 3; Length 404;

Best Local Similarity 98.0%; Pred. No. 2.3e-210;

Matches 386; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 MTDDMLHWKMHNGEKDYPFSDAEMTRRQNDVRGMKKNVDAALFTSYHCINYYSGWL 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

DB 1 MTDDMLHWKMHNGEKDYPFSDAEMTRRQNDVRGMKKNVDAALFTSYHCINYYSGWL 60
 QY 61 YCFGRKRYAMVDHNNATTISAGIDGGPWRSPDNTTYTWRQNFYRAARQLTTGAK 120
 DB 61 YCFGRKRYAMVDHNNATTISAGIDGGPWRSPDNTTYTWRQNFYRAARQLTTGAK 120
 QY 121 RIGIEFDHVNLDPRQLEENLPGVEFVDISQPSMMRTIKSLBEOKLIREGARVCVGA 180
 DB 121 RIGIEFDHVNLDPRQLEENLPGVEFVDISQPSMMRTIKSLBEOKLIREGARVCVGA 180
 QY 181 ACAAAIKAGVPEHEVAIATTNAMIIRIAKSPFVELMDTWTWFOGINTDGAHPVTNRI 240
 DB 181 ACAAAIKAGVPEHEVAIATTNAMIIRIAKSPFVELMDTWTWFOGINTDGAHPVTNRI 240
 QY 241 VOSGDILSINTFPMIFGYTALERTLFCDHVDDASLDIWEKNVAVHRRGLELIKPGARCK 300
 DB 241 VOSGDILSINTFPMIFGYTALERTLFCDHVDDASLDIWEKNVAVHRRGLELIKPGARCK 300
 QY 301 DIAIELNEMRYEMDILKTRSPGYGHSFGVLCYHGREAGVELREDIDTELKPGMVSMMP 360
 DB 301 DIAIELNEMRYEMDILKTRSPGYGHSFGVLCYHGREAGVELREDIDTELKPGMVSMMP 360
 QY 361 MWMLPEGMPGAGGYREHDLIVGEDGAENITGPPGPEHNIIRN 404
 DB 361 MWMLPEGMPGAGGYREHDLIVGEDGAENITGPPGPEHNIIRN 404

RESULT 10

ABR43478
 ID ABR43478 standard; protein; 404 AA.

AC ABR43478;

DT 21-JUL-2003 (first entry)

DE Mutant Erwinia creatinase CTqc2 protein SEQ ID NO.18.

KW Erwinia; creatinase; creatine amidinohydrolase; enzyme; EC 3.5.3.3; Erwinia-type creatinase; creatine; creatinase; uraemia; gigantism; chronic nephritis; acute nephritis; tonic muscular dystrophy; mutant.

OS Erwinia sp.

OS Synthetic.

PN EP1298213-A1.

PD 02-APR-2003.

PF 17-SEP-2002; 2002EP-00020793.

PR 20-SEP-2001; 2001EP-00121780.

PA (HOFF) ROCHE DIAGNOSTICS GMBH.

PA (HOFF) HOFFMANN LA ROCHE & CO AG F.

PI Shao Z, Schumuck R, Kratzsch P, Kenkies J, Weisser H;

DR WPI; 2003-383834/37.

DR N-PSDB; ACC69519.

PT New variant of an Erwinia-type creatinase modified relative to a wild-type creatinase having creatinase activity, useful for determining creatinine and/or creatine concentration in a sample.

PS Example 4; Page 35-36; 51pp; English.

CC The present invention describes a variant of an Erwinia-type creatinase (I) modified relative to a wild-type creatinase having creatinase activity. The variant comprises at least one amino acid substitution at a position of the fully defined 404 amino acid sequence given in ABR43467. CC These are selected from N130, M203, I278, I304 and F395. Creatinase has CC the EC number EC 3.5.3.3 and is also known as creatine amidinohydrolase. CC Also described is a reagent (II) for determining creatine comprising the

CC Erwinia-type creatinase variant. The variant is useful for determining
CC creatinine and/or creatine concentration in a sample. Measuring
CC creatinine and creatine are useful for diagnosing uraemia, chronic
CC nephritis, acute nephritis, gigantism, tonic muscular dystrophy and some
CC other related diseases. The mutant enzymes have improved stability, lower
CC conductivity and/or lower Km-values for creatine; they are much better
CC suited to detection methods for creatine. The present sequence represents
CC a mutant Erwinia creatinase from the present invention
XX

XX Sequence 404 AA;

Query Match 94.3%; Score 2078; DB 6; Length 404;
Best Local Similarity 94.1%; Pred. No. 4.2e-203;

Matches 380; Conservative 14; Mismatches 10; Indels 0; Gaps 0;

QY 1 MTDDMLHWKMHNGEKDYSPPSDAEMTRNDVRGMANKNDVALFTSYHCINYSGL 60
DB 1 MTDDMLHWKMHNGEKDYSPPSDAEMTRNDVRGMANKNDVALFTSYHCINYSGL 60
QY 61 YCFGRKGMVIDHNNATTISAGIDGQPMRSPFGDNITTTDRNDNFYAVRQLTTGAK 120
DB 61 YCFGRKGMVIDHNNATTISAGIDGQPMRSPFGDNITTTDRNDNFYAVRQLTTGAK 120
QY 121 RIGIEFDHVNLDFFRQLLEALPGVEFVDISQPSMMKRTKLSLEOKLIREGARVCVGA 180
DB 121 RIGIEFDHVNLDFFRQLLEALPGVEFVDISQPSMMKRTKLSLEOKLIREGARVCVGA 180
QY 181 ACVAATKGVPRHEVAIATTNMAVRIASFPFVELMDTWFOGINTDGAHPVTNRI 240
DB 181 ACVAATKGVPRHEVAIATTNMAVRIASFPFVELMDTWFOGINTDGAHPVTNRI 240
QY 241 VOSGDILSLNTPPMIFGYTTALERTLFCDHVDASLDIWEKNVAHRRGLELIKPARCK 300
DB 241 VOSGDILSLNTPPMIFGYTTALERTLFCDHVDASLDIWEKNVAHRRGLELIKPARCK 300
QY 301 DIAIELNEMRYREMDLLKXRSFGYGSFGVLCHEYGREAGVELREDIDTELKPGMVSMEP 360
DB 301 DIAIELNEMRYREMDLLKXRSFGYGSFGVLCHEYGREAGVELREDIDTELKPGMVSMEP 360
QY 361 MVMLEPGAGAGYREHDILLVGEDGAENITGPPGPEHNITRN 404
DB 361 MVMLEPGAGAGYREHDILLVGEDGAENITGPPGPEHNITRN 404

RESULT 11
ABR43467 standard; protein; 404 AA.

AC ABR43467;

DT 21-JUL-2003 (first entry)

DE Erwinia sp. (DSM 97-934) creatinase protein SEQ ID NO.2.

KW Erwinia: creatinase; creatine amidinohydrolase; enzyme; EC 3.5.3.3;
KM Erwinia-type creatinase; creatine; creatinine; uraemia; gigantism;
KM chronic nephritis; acute nephritis; tonic muscular dystrophy.

OS Erwinia sp.

PN EP1298213-A1.

PD 02-APR-2003.

PP 17-SEP-2002; 2002EP-00020793.

PR 20-SEP-2001; 2001EP-00121780.

PA (HOFF) ROCHE DIAGNOSTICS GMBH.
PA (HOFF) HOFFMANN LA ROCHE & CO AG F.

PI Shao Z, Schmuck R, Kratzsch F, Kenkles J, Welser H;
KW

DR WPI; 2003-383834/37.
DR N-PSDB; ACC69514.

PT New variant of an Erwinia-type creatinase modified relative to a wild-
PT type creatinase having creatinase activity, useful for determining
PT creatinine and/or creatine concentration in a sample.
XX

PS Claim 1; Page 17-18; 51pp; English.

CC The present invention describes a variant of an Erwinia-type creatinase
CC (1) modified relative to a wild-type creatinase having creatinase
CC activity. The variant comprises at least one amino acid substitution at a
CC position of the fully defined 404 amino acid sequence given in ABR43467;
CC these are selected from N130, M203, I278, I1304 and P395. Creatinase has
CC the EC number EC 3.5.3.3 and is also known as creatine amidinohydrolase.
CC Also described is a reagent (iii) for determining creatine comprising the
CC Erwinia-type creatinase variant. The variant is useful for determining
CC creatinine and/or creatine concentration in a sample. Measuring
CC creatinine and creatine are useful for diagnosing uraemia, chronic
CC nephritis, acute nephritis, gigantism, tonic muscular dystrophy and some
CC other related diseases. The mutant enzymes have improved stability, lower
CC conductivity and/or lower Km-values for creatine; they are much better
CC suited to detection methods for creatine. The present sequence represents
CC Erwinia sp. (DSM 97-934) wild-type creatinase from the present invention
XX

XX Sequence 404 AA;

Query Match 94.9%; Score 2078; DB 6; Length 404;
Best Local Similarity 94.3%; Pred. No. 4.2e-203;

Matches 381; Conservative 14; Mismatches 9; Indels 0; Gaps 0;

QY 1 MTDDMLHWKMHNGEKDYSPPSDAEMTRNDVRGMANKNDVALFTSYHCINYSGL 60
DB 1 MTDDMLHWKMHNGEKDYSPPSDAEMTRNDVRGMANKNDVALFTSYHCINYSGL 60
QY 61 YCFGRKGMVIDHNNATTISAGIDGQPMRSPFGDNITTTDRNDNFYAVRQLTTGAK 120
DB 61 YCFGRKGMVIDHNNATTISAGIDGQPMRSPFGDNITTTDRNDNFYAVRQLTTGAK 120
QY 121 RIGIEFDHVNLDFFRQLLEALPGVEFVDISQPSMMKRTKLSLEOKLIREGARVCVGA 180
DB 121 RIGIEFDHVNLDFFRQLLEALPGVEFVDISQPSMMKRTKLSLEOKLIREGARVCVGA 180
QY 181 ACVAATKGVPRHEVAIATTNMAVRIASFPFVELMDTWFOGINTDGAHPVTNRI 240
DB 181 ACVAATKGVPRHEVAIATTNMAVRIASFPFVELMDTWFOGINTDGAHPVTNRI 240
QY 241 VOSGDILSLNTPPMIFGYTTALERTLFCDHVDASLDIWEKNVAHRRGLELIKPARCK 300
DB 241 VOSGDILSLNTPPMIFGYTTALERTLFCDHVDASLDIWEKNVAHRRGLELIKPARCK 300
QY 301 DIAIELNEMRYREMDLLKXRSFGYGSFGVLCHEYGREAGVELREDIDTELKPGMVSMEP 360
DB 301 DIAIELNEMRYREMDLLKXRSFGYGSFGVLCHEYGREAGVELREDIDTELKPGMVSMEP 360
QY 361 MVMLEPGAGAGYREHDILLVGEDGAENITGPPGPEHNITRN 404
DB 361 MVMLEPGAGAGYREHDILLVGEDGAENITGPPGPEHNITRN 404

RESULT 12

ABR43480 standard; protein; 404 AA.

AC ABR43480;

DT 21-JUL-2003 (first entry)

DE Mutant Erwinia creatinase Ctsd7 protein SEQ ID NO.22.

KW Erwinia: creatinase; creatine amidinohydrolase; enzyme; EC 3.5.3.3;
KW Erwinia-type creatinase; creatine; creatinine; uraemia; gigantism;
KW chronic nephritis; acute nephritis; tonic muscular dystrophy; mutant.


```

XX Erwinia sp.
OS Synthetic.
XX EPI298213-A1.
XX 02-APR-2003.
XX 17-SEP-2002; 2002EP-00020793.
XX 20-SEP-2001; 2001EP-00121780.
XX (HOPE ) ROCHE DIAGNOSTICS GMBH.
XX (HOFF ) HOFFMANN LA ROCHE & CO AG F.
XX Shao Z, Schmuck R, Kratzsch P, Kenklies J, Weisser H;
XX WPI; 2003-383834/37.
XX N-PSDB; ACC69521.
XX
XX New variant of an Erwinia-type creatinase modified relative to a wild-
XX type creatinase having creatinase activity, useful for determining
XX creatinase and/or creatine concentration in a sample.
XX
XX Example 4; Page 42-44; 51pp; English.
XX
XX The present invention describes a variant of an Erwinia-type creatinase
XX (I) modified relative to a wild-type creatinase having creatinase
XX activity. The variant comprises at least one amino acid substitution at a
XX position of the fully defined 404 amino acid sequence given in ABR43467:
XX these are selected from N130, M203, I278, I1304 and P395. Creatinase has
XX the EC number EC 3.5.3.3 and is also known as creatine amidinohydrolase.
XX Also described is a reagent (II) for determining creatine comprising the
XX Erwinia-type creatinase variant. The variant is useful for determining
XX creatinine and/or creatine concentration in a sample. Measuring
XX creatinine and creatine are useful for diagnosing uraemia, chronic
XX nephritis, acute nephritis, gigantism, tonic muscular dystrophy and some
XX other related diseases. The mutant enzymes have improved stability, lower
XX conductivity and/or lower Km-values for creatine. They are much better
XX suited to detection methods for creatine. The present sequence represents
XX a mutant Erwinia creatinase from the present invention
XX
XX Sequence 404 AA;
SQ
Query Match          94.4%; Score 2068; DB 6; Length 404;
Best Local Similarity 93.6%; Pred. No. 4.4e-202;
Matches 378; Conservative 15; Mismatches 11; Indels 0; Gaps 0;
QY 1 MTDDMLHWKMHNGEKDYSPFSDAEMTRQSDVRGMMAKNVDALFTSYHCINYYSGWL 60
DB 1 MTDDMLHWKMHNGEKDYSPFSDAEMTRQSDVRGMMAKNVDALFTSYHCINYYSGWL 60
QY 61 YCYFGRKGYMVIDHNNATTISAGIDGGOPWRRSFGDNITTYDMRRDNFRAVROLTTGAK 120
DB 61 YCYFGRKGYMVIDHNNATTISAGIDGGOPWRRSFGDNITTYDMRRDNFRAVROLTTGAK 120
QY 121 RIGIEFDHVDLDFRRLTEBALPGVEFVDIGQSPMMWRITKLSBEOKLIEGARICDVGGA 180
DB 121 RIGIEFDHVDLDFRRLTEBALPGVEFVDIGQSPMMWRITKLSBEOKLIEGARICDVGGA 180
QY 181 ACAAALIKAGVPEHEVAIATTNMIRIAPSPFVELMDTWTFQSGINTDGAHPVTNRI 240
DB 181 ACAAALIKAGVPEHEVAIATTNMIRIAPSPFVELMDTWTFQSGINTDGAHPVTNRI 240
QY 241 VOSGILSLNTFPMIFGYTALERTLFCDHVDASLIDIEKVAIVARRGLEILIKGARCK 300
DB 241 VOSGILSLNTFPMIFGYTALERTLFCDHVDASLIDIEKVAIVARRGLEILIKGARCK 300
QY 301 DIAIENMYREMDLLIKYRSFGYSHSGVYLCHYGRAGAEVLEIDITELKFGMYVSWEP 360
DB 301 DIAIENMYREMDLLIKYRSFGYSHSGVYLCHYGRAGAEVLEIDITELKFGMYVSWEP 360
QY 361 MVMLEBGMFGAGYREHDLILVIGEDGAENITGPPPEHNITRN 404

```

```

DB 361 MVMLEBGMFGAGYREHDLILVIGEDGAENITGPPPEHNITRN 404
RESULT 13
ABR43474
ID ABR43474 standard; protein; 404 AA.
XX
XX ABR43474;
XX
XX 21-JUL-2003 (first entry)
XX
XX Mutant Erwinia creatinase CTIm24 protein SEQ ID NO:10.
XX
XX Erwinia; creatinase; creatine amidinohydrolase; enzyme; EC 3.5.3.3;
XX Erwinia-type creatinase; creatine; creatinine; uremia; gigantism;
XX chronic nephritis; acute nephritis; tonic muscular dystrophy; mutant.
XX
XX Erwinia sp.
XX Synthetic.
XX
XX EPI298213-A1.
XX 02-APR-2003.
XX 17-SEP-2002; 2002EP-00020793.
XX 20-SEP-2001; 2001EP-00121780.
XX (HOPE ) ROCHE DIAGNOSTICS GMBH.
XX (HOFF ) HOFFMANN LA ROCHE & CO AG F.
XX Shao Z, Schmuck R, Kratzsch P, Kenklies J, Weisser H;
XX WPI; 2003-383834/37.
XX N-PSDB; ACC69515.
XX
XX New variant of an Erwinia-type creatinase modified relative to a wild-
XX type creatinase having creatinase activity, useful for determining
XX creatinase and/or creatine concentration in a sample.
XX
XX Example 4; Page 22-23; 51pp; English.
XX
XX The present invention describes a variant of an Erwinia-type creatinase
XX (I) modified relative to a wild-type creatinase having creatinase
XX activity. The variant comprises at least one amino acid substitution at a
XX position of the fully defined 404 amino acid sequence given in ABR43467:
XX these are selected from N130, M203, I278, I1304 and P395. Creatinase has
XX the EC number EC 3.5.3.3 and is also known as creatine amidinohydrolase.
XX Also described is a reagent (II) for determining creatine comprising the
XX Erwinia-type creatinase variant. The variant is useful for determining
XX creatinine and/or creatine concentration in a sample. Measuring
XX creatinine and creatine are useful for diagnosing uraemia, chronic
XX nephritis, acute nephritis, gigantism, tonic muscular dystrophy and some
XX other related diseases. The mutant enzymes have improved stability, lower
XX conductivity and/or lower Km-values for creatine. They are much better
XX suited to detection methods for creatine. The present sequence represents
XX a mutant Erwinia creatinase from the present invention
XX
XX Sequence 404 AA;
SQ
Query Match          94.4%; Score 2068; DB 6; Length 404;
Best Local Similarity 93.8%; Pred. No. 4.4e-202;
Matches 379; Conservative 15; Mismatches 10; Indels 0; Gaps 0;
QY 1 MTDDMLHWKMHNGEKDYSPFSDAEMTRQSDVRGMMAKNVDALFTSYHCINYYSGWL 60
DB 1 MTDDMLHWKMHNGEKDYSPFSDAEMTRQSDVRGMMAKNVDALFTSYHCINYYSGWL 60
QY 61 YCYFGRKGYMVIDHNNATTISAGIDGGOPWRRSFGDNITTYDMRRDNFRAVROLTTGAK 120
DB 61 YCYFGRKGYMVIDHNNATTISAGIDGGOPWRRSFGDNITTYDMRRDNFRAVROLTTGAK 120

```

QY 121 RIGIEFDHVNLDPRROLEBALPGVEFVDISQPSMMNRITKSLSEOKLIREGARVCDVGGA 180
 DB 121 RIGIEFDHVNLDPRRTLEBALPGVEFVDIGQPSMMNRITKSLSEOKLIREGARICDVGGA 180
 QY 181 ACAAIKAGVPEHEVAIATTNAMIKEIASPPFVELMDWTWTFOSGINTDGAHPVTNRI 240
 DB 181 ACVAIVKAGVPEHEVAIATTNAMIKEIASPPFVELMDWTWTFOSGINTDGAHPVTNRI 240
 QY 241 VOSGDIISLNTFPMIFGYTTALERTLPCDHVDASLDIMEKNVAVHRGELIKPGARCK 300
 DB 241 VOSGDIISLNTFPMIFGYTTALERTLPCDHVDASLDIMEKNVAVHRGELIKPGARCK 300
 QY 301 DIAELNEMVREMDLLKTRSPFGYGHSGFVLCCHYGRAGVELREDIDTELKPGMVSMEP 360
 DB 301 DIAELNEMVREMDLLKTRSPFGYGHSGFVLCCHYGRAGVELREDIDTELKPGMVSMEP 360
 QY 361 MVMLEPGMAGAGYREHDLIVGEDGAENITGPPGPEHNIIIRN 404
 DB 361 MVMLEPGMAGAGYREHDLIVGEDGAENITGPPGPEHNIIIRN 404

RESULT 14

ABR43477
 ID ABR43477 standard; protein; 404 AA.

AC ABR43477;
 XX

DT 21-JUL-2003 (first entry)
 XX

DE Mutant Erwinia creatinase CT2m28 protein SEQ ID NO:16.
 XX

KW Erwinia; creatinase; creatine amidinohydrolase; enzyme; EC 3.5.3.3;
 KM Erwinia-type creatinase; creatine; creatinine; ureaemia; gigantism;
 XX chronic nephritis; acute nephritis; tonic muscular dystrophy; mutant.

OS Erwinia sp.
 OS Synthetic.

XX
 PN EPI298213-A1.

XX
 PD 02-APR-2003.
 XX

PF 17-SEP-2002; 2002EP-00020793.
 XX

PR 20-SEP-2001; 2001EP-00121780.
 XX

PA (HOFF) ROCHE DIAGNOSTICS GMBH.
 PA (HOFF) HOFFMANN LA ROCHE & CO AG F.

XX
 PI Shao Z, Schmuck R, Kratzsch P, Kenkiles J, Weisser H;
 XX

DR WPI; 2003-383834/37.
 DR N-PSDB; ACC69518.

XX
 PS Example 4; Page 32-33; 51pp; English.

XX The present invention describes a variant of an Erwinia-type creatinase
 CC (i) modified relative to a wild-type creatinase having creatinase
 CC activity. The variant comprises at least one amino acid substitution at a
 CC position of the fully defined 404 amino acid sequence given in ABR43467;
 CC these are selected from N130, M203, I278, I1304 and F395. Creatinase has
 CC the EC number EC 3.5.3.3 and is also known as creatine amidinohydrolase.
 CC Also described is a reagent (ii) for determining creatine creatinizing the
 CC Erwinia-type creatinase variant. The variant is useful for determining
 CC creatinine and/or creatine concentration in a sample. Measuring
 CC creatinine and creatine are useful for diagnosing uremia, chronic
 CC nephritis, acute nephritis, gigantism, tonic muscular dystrophy and some
 CC other related diseases. The mutant enzymes have improved stability, lower
 CC conductivity and/or lower Km-values for creatine: they are much better

CC suited to detection methods for creatine. The present sequence represents
 CC a mutant Erwinia creatinase from the present invention

XX
 SQ Sequence 404 AA;

Query Match 94.3%; Score 2066; DB 6; Length 404;
 Best Local Similarity 93.6%; Pred. No. 7e-202;
 Matches 378; Conservative 16; Mismatches 10; Indels 0; Gaps 0;

QY 1 MTDDMLHKMKHNGEKDVSPPSDAEMTRQNDVRCGMKANNVDALFTSYHCINTYSGWL 60
 DB 1 MTDDMLHKMKHNGEKDVSPPSDAEMTRQNDVRCGMKANNVDALFTSYHCINTYSGWL 60
 QY 61 YCFGRKGYMIDHNNATTISAGIDGQPMRSPFSDNITTYDMRDNDFRAVROLTTGAK 120
 DB 61 YCFGRKGYMIDHNNATTISAGIDGQPMRSPFSDNITTYDMRDNDFRAVROLTTGAK 120
 QY 121 RIGIEFDHVNLDPRROLEBALPGVEFVDISQPSMMNRITKSLSEOKLIREGARVCDVGGA 180
 DB 121 RIGIEFDHVNLDPRRTLEBALPGVEFVDIGQPSMMNRITKSLSEOKLIREGARICDVGGA 180
 QY 181 ACAAIKAGVPEHEVAIATTNAMIKEIASPPFVELMDWTWTFOSGINTDGAHPVTNRI 240
 DB 181 ACVAIVKAGVPEHEVAIATTNAMIKEIASPPFVELMDWTWTFOSGINTDGAHPVTNRI 240
 QY 241 VOSGDIISLNTFPMIFGYTTALERTLPCDHVDASLDIMEKNVAVHRGELIKPGARCK 300
 DB 241 VOSGDIISLNTFPMIFGYTTALERTLPCDHVDASLDIMEKNVAVHRGELIKPGARCK 300
 QY 301 DIAELNEMVREMDLLKTRSPFGYGHSGFVLCCHYGRAGVELREDIDTELKPGMVSMEP 360
 DB 301 DIAELNEMVREMDLLKTRSPFGYGHSGFVLCCHYGRAGVELREDIDTELKPGMVSMEP 360
 QY 361 MVMLEPGMAGAGYREHDLIVGEDGAENITGPPGPEHNIIIRN 404
 DB 361 MVMLEPGMAGAGYREHDLIVGEDGAENITGPPGPEHNIIIRN 404

RESULT 15

ABR43475
 ID ABR43475 standard; protein; 404 AA.

AC ABR43475;
 XX

DT 21-JUL-2003 (first entry)
 XX

DE Mutant Erwinia creatinase CT2m9 protein SEQ ID NO:12.
 XX

KW Erwinia; creatinase; creatine amidinohydrolase; enzyme; EC 3.5.3.3;
 KM Erwinia-type creatinase; creatine; creatinine; ureaemia; gigantism;
 XX chronic nephritis; acute nephritis; tonic muscular dystrophy; mutant.

OS Erwinia sp.
 OS Synthetic.

XX
 PN EPI298213-A1.

XX
 PD 02-APR-2003.
 XX

PF 17-SEP-2002; 2002EP-00020793.
 XX

PR 20-SEP-2001; 2001EP-00121780.
 XX

PA (HOFF) ROCHE DIAGNOSTICS GMBH.
 PA (HOFF) HOFFMANN LA ROCHE & CO AG F.

XX
 PI Shao Z, Schmuck R, Kratzsch P, Kenkiles J, Weisser H;
 XX

DR WPI; 2003-383834/37.
 DR N-PSDB; ACC69518.

XX
 PS Example 4; Page 32-33; 51pp; English.

XX The present invention describes a variant of an Erwinia-type creatinase
 CC (i) modified relative to a wild-type creatinase having creatinase
 CC activity. The variant comprises at least one amino acid substitution at a
 CC position of the fully defined 404 amino acid sequence given in ABR43467;
 CC these are selected from N130, M203, I278, I1304 and F395. Creatinase has
 CC the EC number EC 3.5.3.3 and is also known as creatine amidinohydrolase.
 CC Also described is a reagent (ii) for determining creatine creatinizing the
 CC Erwinia-type creatinase variant. The variant is useful for determining
 CC creatinine and/or creatine concentration in a sample. Measuring
 CC creatinine and creatine are useful for diagnosing uremia, chronic
 CC nephritis, acute nephritis, gigantism, tonic muscular dystrophy and some
 CC other related diseases. The mutant enzymes have improved stability, lower
 CC conductivity and/or lower Km-values for creatine: they are much better

PT New variant of an Erwinia-type creatinase modified relative to a wild-
 type creatinase having creatinase activity, useful for determining

PT creatinine and/or creatine concentration in a sample.

XX Example 4, Page 25-26; 51pp; English.

XX The present invention describes a variant of an Erwinia-type creatinase
CC (I) modified relative to a wild-type creatinase having creatinase
CC activity. The variant comprises at least one amino acid substitution at a
CC position of the fully defined 404 amino acid sequence given in ABR43467;
CC these are selected from N130, M203, I278, I1304 and F395. Creatinase has
CC the EC number EC 3.5.3.3 and is also known as creatine amidinohydrolase.
CC Also described is a reagent (II) for determining creatine comprising the
CC Erwinia-type creatinase variant. The variant is useful for determining
CC creatinine and/or creatine concentration in a sample. Measuring
CC creatinine and creatine are useful for diagnosing uraemia, chronic
CC nephritis, acute nephritis, gigantism, tonic muscular dystrophy and some
CC other related diseases. The mutant enzymes have improved stability, lower
CC conductivity and/or lower Km-values for creatine; they are much better
CC suited to detection methods for creatine. The present sequence represents
CC a mutant Erwinia creatinase from the present invention

XX Sequence 404 AA;

Query Match 94.2%; Score 2064; DB 6; Length 404;

Best Local Similarity 93.6%; Pred. No. 1.1e-201;

Matches 378; Conservative 16; Mismatches 10; Indels 0; Gaps 0;

QY 1 MTDDMLHYMKWNGEKDYSPFSDAEMTRONDVKGMAKNNDALFTSYHCINITYSGML 60
DB 1 MTDDMLHYMKWNGEKDYSPFSDAEMTRONDVKGMAKNNDALFTSYHCINITYSGFL 60
QY 61 YCYFGRKKGWYIDHNNATTISAGIDGQPMWRSPGDNITTYTDMRRDNFYRAVROLTTGAK 120
DB 61 YCYFGRKKGWYIDQDHATTISAGIDGQPMWRSPGDNITTYTDMRRDNFYQAVROLTPGAR 120
QY 121 RIGIEFDHVNLDPRQLEBALPGVEFVDISQPMWRRTIKSLSEQKLIREGARVCVGA 180
DB 121 RIGIEFDHVNLDPRQLEBALPGVEFVDISQPMWRRTIKSLSEQKLIREGARICDVGA 180
QY 181 ACAAAIKAGVPEHEVAIATNNMIREIAKSPPELMDTWTFQSGINTDGAHPVTNRI 240
DB 181 ACVAAVKAGVPEHEVAIATNNAVREIAKSPPELMDTWTFQSGINTDGAHPVTNRI 240
QY 241 VOSGDILSLNFPMTFGYTTALERTLFCDHVDASLDIWEKNVAVHRGLLEIKPGARCK 300
DB 241 VOSGDILSLNFPMTFGYTTALERTLFCDHVDASLDIWEKNVAVHRGLLEIKPGARCK 300
QY 301 DIAIELNEMYREWDLKTRSPFGYGHSPGYLCHYGREAGVELREDIDTELKPMVVSMEP 360
DB 301 DIAIELNEMYREWDLKTRSPFGYGHSPGYLCHYGREAGVELREDIDTELKPMVVSMEP 360
QY 361 MVMLEPGMFGAGYREHDILYVGEDGAENITGPPGPEHNITRN 404
DB 361 MVMLEPGMFGAGYREHDILYVGEDGAENITGPPGPEHNITRN 404

Search completed: July 8, 2005, 01:34:11

Job time : 84 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 8, 2005, 01:01:21 ; Search time 24 Seconds
(without alignments)
1619.649 Million cell updates/sec

Title: US-10-807-228a-1

Perfect score: 2190
Sequence: 1 MTDDMLHVMKMHNGEKDYP.....DGAENITGPFPGHEHNIIRN 404

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Query length	DB ID	Description
1	1349	61.6	411	2	I39809 creatinase (EC 3.5
2	1329.5	60.7	411	2	T44250 creatinase (EC 3.5
3	1048.5	47.9	378	2	UH0134 creatinase (EC 3.5
4	257	11.7	363	1	G69869 Xaa-Pro dipeptidase
5	255	11.6	353	1	C69860 Xaa-Pro dipeptidase
6	248.5	11.3	351	2	E75088 cobalt-dependent p
7	247	11.3	355	2	H83999 Xaa-Pro dipeptidase
8	246.5	11.3	351	1	G71056 Probable X-Pro dip
9	232	10.6	406	2	C83867 Xaa-Pro dipeptidase
10	231.5	10.6	365	2	AD1634 X-Pro dipeptidase
11	230.5	10.5	365	2	AB1272 X-Pro dipeptidase
12	229	10.5	361	1	PF5012 hypothetical prote
13	229	10.5	361	2	A91037 probable peptidase
14	229	10.5	361	2	C85881 probable peptidase
15	227.5	10.4	364	2	C84047 prolidase (proline
16	226.5	10.3	348	2	T46973 X-Pro dipeptidase
17	223.5	10.2	353	2	AB1244 aminopeptidase p h
18	223.5	10.2	353	2	AF1606 aminopeptidase p h
19	222.5	10.2	376	2	A86974 Probable cytoplasm
20	222	10.1	356	2	C97158 aminopeptidase p A
21	219.5	10.0	349	2	D75419 proline dipeptidase
22	217	9.9	356	2	F81657 proline dipeptidase
23	216	9.9	372	1	C70658 probable pepo - My
24	215.5	9.8	352	2	C86711 aminopeptidase p l
25	210	9.6	376	2	T35868 probable dipeptida
26	208.5	9.5	393	2	D95894 probable hydrolase
27	203	9.0	353	2	A89933 Xaa-Pro dipeptidase
28	197.5	9.0	348	2	H90572 xaa-pro aminopepti
29	197	9.0	356	2	H71496 probable aminopept

30	193.5	8.8	349	2	A72750 probable Xaa-Pro d
31	192.5	8.8	356	1	A71089 probable dipeptida
32	191.5	8.7	359	2	G72425 hypothetical prote
33	191	8.7	362	2	A86831 X-Pro dipeptidase
34	190.5	8.7	360	2	D95185 proline dipeptidase
35	190.5	8.7	392	2	A13141 Xaa-Pro dipeptidase
36	190.5	8.7	392	2	C95146 Xaa-Pro dipeptidase
37	189.5	8.7	360	2	D98052 X-Pro dipeptidase
38	189	8.6	353	2	D97893 X-Pro dipeptidase
39	185.5	8.5	358	2	A97243 Xaa-Pro aminopepti
40	183	8.4	353	2	G95021 peptidase M24 faml
41	182	8.3	358	2	H89954 hypothetical prote
42	181.5	8.3	434	2	C95362 probable aminopept
43	180.5	8.2	368	1	S52302 X-Pro dipeptidase
44	175.5	8.0	354	1	C70433 probable X-Pro dip
45	175.5	8.0	355	2	C75134 x-pro aminopeptida

ALIGNMENTS

RESULT 1

139809 creatinase (EC 3.5.3.3) - Bacillus sp.
C/Species: Bacillus sp.
C/Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #ext_change 09-Jul-2004
C/Accession: I39809; I39976
R/Suzuki, K.; Sagai, H.; Sugiyama, M.; Imamura, S.
J. Ferment. Bioeng. 76, 77-81, 1993
J. Ferment. Bioeng. 77, 231-234, 1994
A/Title: Molecular cloning and high expression of the Bacillus creatinase gene in Esch
A/Reference number: I39809
A/Accession: I39809
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-411 <RES>
A/Cross-references: UNIPROT:P38487; GB:D14463; NID:G500610; PIDN:BA03358.1; PID:G5006
R/Suzuki, K.; Sagai, H.; Imamura, S.; Sugiyama, M.
J. Ferment. Bioeng. 77, 231-234, 1994
A/Title: Cloning, sequencing, overexpression in Escherichia coli of a sarcosine oxidase
A/Reference number: I39976
A/Accession: I39976
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-46 <RES>
A/Cross-references: GB:D16521; NID:G984787; PIDN:BA03968.1; PID:G840668
C/Genetics:
A/Gene: cre
C/Superfamily: X-Pro aminopeptidase
C/Keywords: hydrolase

Query Match 61.6%; Score 1349; DB 2; Length 411;
Best Local Similarity 63.9%; Pred. No. 1.9e-106;
Matches 253; Conservative 49; Mismatches 90; Indels 4; Gaps 1;

QY	12	HNHGDYSPSPDAEMTRQNDVGMMAKNNVAALPTSHCHINYSGLYCYFGRXYGNV	71
DB	15	NGSEKVPFTFSKXEMTRNTRLRBYAKAGIDAVMTSHNINYSDFLYTSFNRSAIV	74
QY	72	IDHNNATTISAGIDGQFWRSGDNITTYDRRRNRYAVALTT----GAKRIGIEFD	127
DB	75	VYQDGHVTVSANIDGMPWRSGFDENIYTTDKRKNFLYAVKVNIEGFSGRCLGVND	134
QY	128	HYNDPFRQLEALPGVEFVDISQPSMMRTIKSLBEOKLREGARVCDVGAAACAAIK	187
DB	135	HMTLDLRQVODALPNTLTVVVSQAVMGHRMFKSDEIDLLINGCARIDIGAAVVEAIR	194
QY	168	AGVPEHYVALNTNMMIREIKSPFVELMDTWTFQSGINNDGANNPTVNTNIVYSGDIL	247
DB	195	BSVPBYEVALHGTBMVREIATYPAHELTDWIWTFQSGINTDGAHNAATSKLDRGDIL	254
QY	248	SLATPMTIFGYVTLALERTLPCDHVDASLDIWEKVAAYHRRGLIKFGARCKDIAIELN	307
DB	255	SLNCPMTIAGYVTLALERTLFLBEVSDRLLEMEINCKHRRGLIKFGARCKDIAIELN	314

QY 308 EMTRENDLKYRSFGYSHSGYVLTCHYGRAGVLEARDIDTELKPGWVMSMEPMMLPEG 367
 DB 315 EIRREHDLNKRFTGYSFGVSHHYGRAGLELRIDETVLEPGWVMSMEPMMLPEG 374
 QY 368 MPGAGYREHDLIVGEGAGENTGPPGPEPHNIIR 403
 DB 375 BPGAGYREHDLIVSNGENTGPPGPEPHNIIR 410

RESULT 2

T44250
 creatinase (EC 3.5.3.3) [validated] - *Archrobacter* sp. (strain TE1826)

C/Species: *Archrobacter* sp.

A/Variety: strain TE1826

C/Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 12-Jul-2004

C/Accession: T44250

R/Nishiya, Y.; Toda, A.; Imanaka, T.

Mol. Gen. Genet. 257, 581-586, 1998

A/Title: Gene cluster for creatinase degradation in *Archrobacter* sp. TE1826.

A/Reference number: 222735; MUID:9822334; PMID:9563845

A/Accession: T44250

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-411 <NIS>

A/Cross-reference: EMBL:AB007122; PDB:BA25928.1

A/Experimental source: strain TE1826

C/Genetics:

A/Function: crea

A/Description: EC 3.5.3.3 [validated, MUID:9822334]

C/Keywords: hydrolase

Query Match 60.7%; Score 1329.5; DB 2; Length 411;
 Best Local Similarity 63.0%; Pred. No. 8.4e-105;
 Matches 250; Conservative 55; Mismatches 87; Indels 5; Gaps 3;

QY 12 HNGEKYSPSPDAEMTRQNDVGRMAKNNVDAALFTSYHCINYSGLCYRGRKXGMV 71
 DB 14 HNGEKYSPSPDAEMTRQNDVGRMAKNNVDAALFTSYHCINYSGLCYRGRKXGMV 73
 QY 72 IDHNNATTISAGIDGQPMRRSFGDNTTYTDMRDNFRAVROL--TTGAK--RIGIEFD 127
 DB 74 VTQKAYTVSANIDGMPRRSYDENIVYTDKRDNYFALQKLEAGVKKALGIEED 133
 QY 128 HYNLDFRRQLEALPGVEFVDISQPSMMRTIKSLSEOKLIREGARVCDDGAAALAAIK 187
 DB 134 HVSIDLKRFSDPFNFELVHSQDVMKQRMKSAEIRHKNGARLADIGYAVVEAIO 193
 QY 188 AGPHEHVAATTAATMARELAKSPPELMDTWTWFOGINTDGAHPVTNRIVOSGDI 247
 DB 194 EGVEYEVALLAGSVMARELAKLPOSELRTWTWFOGINTDGAHSMATSKVQKEIL 253
 QY 248 SLNTPFMIFGYTALERTLFCDDYDASLDIWEKNVAVHRRGLIRKARCKDIAELN 307
 DB 254 SLNTPFMIFGYTALERTLFCDDYDASLDIWEKNVAVHRRGLIRKARCKDIAELN 313
 QY 308 EMTRENDLKYRSFGYSHSGYVLTCHYGRAGVLEARDIDTELKPGWVMSMEPMMLPEG 367
 DB 315 EIRREHDLNKRFTGYSFGVSHHYGRAGLELRIDETVLEPGWVMSMEPMMLPEG 374
 QY 368 MPGAGYREHDLIVGEGAGENTGPPGPEPHNIIR 403
 DB 375 BPGAGYREHDLIVSNGENTGPPGPEPHNIIR 410

RESULT 3

JH0134
 creatinase (EC 3.5.3.3) - *Flavobacterium* sp.

N/Alternate names: creatine amidohydrolase

C/Species: *Flavobacterium* sp.

C/Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 09-Jul-2004

C/Accession: JH0134

R/Koyama, Y.; Kitao, S.; Yamamoto-Otake, H.; Suzuki, M.; Nakano, E.
 Agric. Biol. Chem. 54, 1453-1457, 1990
 A/Title: Cloning and expression of the creatinase gene from *Flavobacterium* sp. U-188
 A/Reference number: JH0134; MUID:91103958; PMID:1368564
 A/Accession: JH0134
 A/Molecule type: DNA
 A/Residues: 1-378 <KOY>
 A/Cross-reference: UNIPROT:P19213
 A/Experimental source: strain U-188
 C/Comment: This enzyme catalyzes the hydrolysis of creatine to sarcosine and urea.
 C/Superfamily: X-Pro aminopeptidase
 C/Keywords: hydrolase

Query Match 47.9%; Score 1048.5; DB 2; Length 378;
 Best Local Similarity 59.9%; Pred. No. 5e-81;
 Matches 200; Conservative 42; Mismatches 91; Indels 1; Gaps 1;

QY 5 MTHVMKHNKEDYSPSPDAEMTRQNDVGRMAKNNVDAALFTSYHCINYSGLCYR 64
 DB 3 MPTLIRNGEKYSGTFSAGVYANRAKLRHLEAENIDAALFTSYHCINYSGLCYR 62
 QY 65 GRKRYGVIDHNNATTISAGIDGQPMRRSFG--DNITYTDMRDNFRAVROLTTGARRIG 123
 DB 63 GRPYALVTVDDVVISANIDGQPMRRRTVGTNDIVYTDKRDNYFALQKLEAGVKKALGIEED 122
 QY 124 IERDYNLDFRRQLEALPGVEFVDISQPSMMRTIKSLSEOKLIREGARVCDDGAAACA 183
 DB 123 IERDYNLDFRRQLEALPGVEFVDISQPSMMRTIKSLSEOKLIREGARVCDDGAAACA 182
 QY 184 AAKAGVPEHEVAATTAATMARELAKSPPELMDTWTWFOGINTDGAHPVTNRIVOS 243
 DB 183 EALRDQVEYEVALLAGSVMARELAKLPOSELRTWTWFOGINTDGAHSMATSKVQKEIL 242
 QY 244 GDLISLTPFMIFGYTALERTLFCDDYDASLDIWEKNVAVHRRGLIRKARCKDIA 303
 DB 243 GDLISLTPFMIFGYTALERTLFCDDYDASLDIWEKNVAVHRRGLIRKARCKDIA 302
 QY 304 IETNEMREHDLIVGEGAGENTGPPGPEPHNIIR 403
 DB 303 KEINERFLKRDILQYRTFGYSHSGYVLTCHYGRAGVLEARDIDTELKPGWVMSMEPMMLPEG 367

RESULT 4

G69869
 Xaa-Pro dipeptidase homolog ykvy - *Bacillus subtilis*

C/Species: *Bacillus subtilis*

C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004

C/Accession: G69869

R/Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bert
 C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; C
 A.; Ehrlich, S.D.; Emmerson, P.T.; Estlin, K.D.; Ewington, J.; Fabret, C.; Ferrati, E
 Nature 390, 249-256, 1997

A/Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall
 lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holzapfel, S.; Hosono, S.; Hullo, M.
 Koeter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino
 A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue
 Y., M.; Ogawa, K.; Ogiwara, A.; Oudeg, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetel
 Rieger, M.; Rivolta, C.; Rocha, E.; Rothe, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanio
 A.; Authors: Schleich, S.; Schroeter, R.; Scrofano, F.; Sekiguchi, J.; Sekowska, A.; Ser
 akeuchi, M.; Yamashita, A.; Yamano, T.; Teipstra, P.; Toynon, A.; Tosato, V.; Uchiyama
 T.; Wintner, P.; Wipac, A.; Yamano, T.; Yamano, T.; Yamano, T.; Yamano, T.; Yamano, T.
 A.; Authors: Yoshikawa, H.F.; Zumschein, E.; Yoshikawa, H.; Danchin, A.

A/Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*

A/Reference number: A69580; MUID:98044033; PMID:9384377

A/Accession: G69869

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Cross-reference: UNIPROT:Q31689; GB:Z99111; GB:AL009126; NID:92633699; PDB:1AB125

A/Experimental source: strain 168

C/Genetics:

A/Function: ykvy

C/Superfamily: X-Pro aminopeptidase

Qy 388 ENT 391
Db 341 RRLT 344

RESULT 7

Xaa-Pro dipeptidase BH2800 (imported) - Bacillus halodurans (strain C-125)
C/Species: Bacillus halodurans
C/Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C/Accession: H83999
R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A/Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A/Reference number: A83650; MUID:20512582; PMID:11058132
A/Accession: H83999
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1355 <STO>
A/Cross-references: UNIPROT:Q9K950; GB:AP001516; GB:BA000004; NID:G10175192; PIDN:BA065
A/Experimental source: strain C-125
C/Genetics:
A:Gene: BH2800
C/Superfamily: X-Pro aminopeptidase

Query Match 11.3%; Score 247; DB 2; Length 355;
Best Local Similarity 27.6%; Pred. No. 4,1e-13;
Matches 78; Conservative 55; Mismatches 122; Indels 28; Gaps 7;

Qy 118 GARRIGEFHVNLDPRROLEALPGVEFVDISQPSMMWRTYKSLEROKLIREGARVCV 177
Db 87 GIGQDGEKSHVTFE-TYELINKLVSELPVAGLVENIRLIKQETELQIQEAMNTADA 145
Qy 178 GGAACAARAKGVEHEVAIATTNAMIETAKSPPELMDTWTFQSGINTGANNPT 237
Db 146 APAMITTYIRAGVTEREVANLEBFRMRKQAGSSFDIV-----ASGRSALPHGVAS 199
Qy 238 NRIVQSGDILSTNFPPIFGYTTALERTLECDHVDASLDIMEKNVAVHRRGLELTPGA 297
Db 200 DKYIEKELVTLDPGAYYKGCSDITRTVAVGINDLRKITYDVTLAOLKMGEMGIRPGI 259
Qy 298 RCDIATLELNMVREMDLKYRSFGYGHSPFVLCHYGRAGVLR-----DIDTELK 351
Db 260 TGR---EADALTRDHTAK---GVEYEG---HSTGHGGLGVHEHGPGLSMKSKAVLK 308
Qy 352 PGWVSMPEWMLPEGMPAGGYREHDLIVGEDGAEINTGFP 394
Db 309 PGWVTVTEPGIYI---SGVGTRIEDDVITTESGNSLTKSP 347

RESULT 8

G71056
Probable X-Pro dipeptidase - Pyrococcus horikoshii
C/Species: Pyrococcus horikoshii
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 12-Jul-2004
C/Accession: G71056
R/Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekit
M.; Ofuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi
DNA Res. 5, 55-76, 1998
A/Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a
A/Reference number: A71000; MUID:98344137; PMID:9679194
A/Accession: G71056
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1351 <KMW>
A/Cross-references: UNIPROT:O58885; GB:AP000005; NID:G3236132; PIDN:BA30249.1; PID:G325
A/Experimental source: strain OT3
A/Note: this accession replaces an interim accession for a sequence replaced by GenBank
C/Genetics:
A:Gene: PH1149

Query Match 11.3%; Score 246.5; DB 1; Length 351;

Best Local Similarity 23.4%; Pred. No. 4.5e-13;
Matches 86; Conservative 75; Mismatches 179; Indels 27; Gaps 8;

Qy 26 MTRRQNDVGRMAKNVDAALFTSYHCINTYSGMLCYGRKKGMVTDHNNATISGID 85
Db 4 MNEKVKKIIIEFMDKNSIDVLIKNPNVYIISGASPLAGY---LITGSATLVYPELE 60
Qy 86 GCGPWRSPGDNNTTYDWR-DNFRVAVQLTGARRIGIEFDHVNLDPRROLEALPGV 144
Db 61 YEMAKES--NIPVEFKKMDFFYOLE---GISLIGIE-SSLPGFIEELKKKANIK 112
Qy 145 EFDVDSQPSMMWRTYKSLEROKLIREGARVCVGAACAARAKGVEHEVAIATTNAMI 204
Db 113 EFKVDVDRDKRIIESEKIKIIEKACEIADKAVMAAIEITEGKEREVAKVEYIMK 172
Qy 205 REIATKSPPELMDTWTFQSGINTGANNPTNRIVQSGDILSTNFPPIFGYTTALER 264
Db 173 MNGAEKPAFD-----TIASGRSALPHGVASDKRIERDVLVIDGALYOHYNSDITR 226
Qy 265 TLECDHVDASLDIMEKNVAVHRRGLELTPGARCDIATLELNMVREMDLKYRSFGY 324
Db 227 TLVVGSPNEKQKEIYIVLEBAQKAVESAKPGITTAELDSIANIITAEYGYEYFNHSLG 286
Qy 325 HSPFVLCHYGRAGVLR-----DIDTELKPGWVSMPEWMLPEGMPAGGYREHDLIVGE 384
Db 287 HGVGLVHEWEPVRSQYD-----ETVLRGVVITIEPGIYIPK---IGVRIEDTILITK 337
Qy 385 DGEANT 391
Db 338 NGSKRLL 344

RESULT 9

C83867
Xaa-Pro dipeptidase BH1739 (imported) - Bacillus halodurans (strain C-125)
C/Species: Bacillus halodurans
C/Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C/Accession: C83867
R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hi
Nucleic Acids Res. 28, 4317-4331, 2000
A/Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans an
A/Reference number: A83650; MUID:20512582; PMID:11058132
A/Accession: C83867
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1406 <STO>
A/Cross-references: UNIPROT:Q9K950; GB:AP001513; GB:BA000004; NID:G10174345; PIDN:BA065
A/Experimental source: strain C-125
C/Genetics:
A:Gene: BH1739

Query Match 10.6%; Score 232; DB 2; Length 406;
Best Local Similarity 24.0%; Pred. No. 9.3e-12;
Matches 96; Conservative 73; Mismatches 179; Indels 52; Gaps 15;

Qy 21 FSDAEMTRRQNDVGRMAKNVDAALF--TSYHCINTYSGMLCYFYGRKKGMVTDHNN 77
Db 2 FSLMEIRYRLNRETKGMDEBIVLVSNPSNMYISGSAMSF-YHQLLVTLDDPOP 60
Qy 78 TTISAGIDGQPWRRSPGDN--ITYTD-----WRDNYRAVRLQTLTGAKRIGIEF 126
Db 61 LWTGREMDASSVYKTTWLBQVTPYRDHYQSETHPMDFVNIKEIQGKRTIGVEM 120
Qy 127 D--HVNLDPRROLEALGVEFVDISQPSMMWRTYKSLEROKLIREGARVCVGAACA 183
Db 121 DAHYFTGLCYQR-LOGLTNGTFKNAVTLINWRLIKSDQELQVWRKAATIAENAKAY 179
Qy 184 AAIKAGVPEHEVAIATTNAMIIR---ETAKSPPELMDTWTFQSGINTGANNPTNRI 240
Db 180 DTYNVGVRECDVAALSHAIKGTAPFGDYSIVM-----LPTGENTSCPHLTWTDRT 234
Qy 241 VQSGDILSTNFPPIFGYTTALERTLFC---DHVDASLDIMEKNVAVHRRGLELTPG 236

Db 235 YOEGBDYLTVELIAGCYKRYHVPARTVSLGVAPEHVKEIAKVIE---GIHET-LQMIRPG 290
 QY 297 ARCDIAIETENMYREMDLTKYRSFGYCHYGRAGVELREDIDTELKPGWV 356
 Db 291 VAAEVAATNQSISKIGFEKNSRLGY--SIGLSFPPMGHWTSLKOGDTLLKPNMTF 348
 QY 357 SMEPMVLPKMGPGAGYREHDI---VGEDAENIT 391
 Db 349 HL-----MPGI-WYEDYGEVETESIRITDDGVLELT 378

RESULT 10

AD1634
 X-Pro dipeptidase homolog jlm1613 [imported] - *Listeria innocua* (strain Clij11262)
 C/Species: *Listeria innocua*
 C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
 C/Accession: AD1634
 R/Glaeser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurge, O.; Entian, K.D.; Fshhi, H.; Science 294, 849-852, 2001
 A/Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapk, G.; Madueno, E.; Maitournam, A.; Mok, C.; Schluter, T.; Simes, N.; Tietz, A.; Vazquez-Boland, J.A.; Voss, H.; Wehler, A./Title: Comparative genomics of *Listeria* species.
 A/Reference number: AB1077; MUID:21537279; PMID:11679669
 A/Accession: AD1634
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-365 <GLA>
 A/Cross-references: UNIPROT:Q92BD7; GB:AL592022; PIDN:CAC96844.1; PID:g16411007; GSPDB:C
 A/Experimental source: strain Clij11262
 C/Genetics:
 A/Gene: jlm1613
 C/Superfamily: X-Pro aminopeptidase

Query Match 10.6%; Score 231.5; DB 2; Length 365;
 Best Local Similarity 22.1%; Pred. No. 8.9e-12;
 Matches 85; Conservative 69; Mismatches 188; Indels 43; Gaps 9;

QY 26 MTRKQNDVRGMKANNVDAALFTSYHCINYSGLYCYFGKRYGMVIDHNNATTISA--- 82
 Db 1 MEKNIDVLQWMLDQGAELVGLFDPENIAVFGSHSEPHRVLGLAVFSDEPLFTPAL 60
 QY 83 ---GIDGQPMRBSFGDNITTYTDMRDNFR---AVROLTTGAKRGIFEDVNDLFR 135
 Db 61 EVDVVGQDMWTHAYGNDT-----ENPKIADETIKKAVANPSKAIKKHMSVRYE 114
 QY 136 QLEBALPGVEFVNDISOPSMRRTIKSLBEOKLIREGARVCDVGAACAALIKAGVPEHEV 195
 Db 115 QLAGLPSGSSFFIIEKHIEIRLIKTEALIKILKALADY-----AVQGV--DEI 165
 QY 196 AIAATTNAMIKEIAKSPF-----VELMDTWTFQSGINTDGAHPVNTNRIVOSGDIISL 249
 Db 166 AEGKTEA---EIVAKIEYEMKKKGVTAHSPDTWVLTKGNALPHGTFGETRIKKGDLVLF 222
 QY 250 NTFPMIFGYTALERTLFCDHVDASLDIWEKNVAVHRLGLIKRGARCKDAIENEM 309
 Db 223 DLGAVHKGYCSDITRIVAFSDITDEOKIYDYLEAOVAADVKKYKIKASEIDLTARNI 282
 QY 310 YREMDLTKRSPFGYCHYGRAGVELREDIDTELKPGWVSMPEMVLPEGM 369
 Db 283 IRRAGGDFPRLHGLGLASVHEF-----PSITETNMELQENMVFTEPGIY----VP 333
 QY 370 GAGGYREHDLIVGEDAENITGFP 394
 Db 334 GVAGVRIEDDLVYTKDGVQLTEFP 358

RESULT 11

AB1272
 X-Pro dipeptidase homolog jmo1578 [imported] - *Listeria monocytogenes* (strain EGD-e)
 C/Species: *Listeria monocytogenes*
 C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004

C/Accession: AB1272
 R/Glaeser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurge, O.; Entian, K.D.; Fshhi, D.; Jones, L.M.; Karst, U.; Science 294, 849-852, 2001
 A/Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapk, G.; Madueno, E.; Maitournam, A.; Mok, C.; Schluter, T.; Simes, N.; Tietz, A.; Vazquez-Boland, J.A.; Voss, H.; Wehler, A./Title: Comparative genomics of *Listeria* species.
 A/Reference number: AB1077; MUID:21537279; PMID:11679669
 A/Accession: AB1272
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-365 <GLA>
 A/Cross-references: UNIPROT:Q8Y6V3; GB:NC_003210; PIDN:CAC99656.1; PID:g16411007; GSPDB:C
 A/Experimental source: strain EGD-e
 C/Genetics:
 A/Gene: jmo1578
 C/Superfamily: X-Pro aminopeptidase

Query Match 10.5%; Score 230.5; DB 2; Length 365;
 Best Local Similarity 22.1%; Pred. No. 1.1e-11;
 Matches 85; Conservative 67; Mismatches 190; Indels 43; Gaps 9;

QY 26 MTRKQNDVRGMKANNVDAALFTSYHCINYSGLYCYFGKRYGMVIDHNNATTISA--- 82
 Db 1 MEKNIDVLQWMLDQGAELVGLFDPENIAVFGSHSEPHRVLGLAVFSDEPLFTPAL 60
 QY 83 ---GIDGQPMRBSFGDNITTYTDMRDNFR---AVROLTTGAKRGIFEDVNDLFR 135
 Db 61 EVDVVGQDMWTHAYGNDT-----ENPKIADETIKKAVANPSKAIKKHMSVRYE 114
 QY 136 QLEBALPGVEFVNDISOPSMRRTIKSLBEOKLIREGARVCDVGAACAALIKAGVPEHEV 195
 Db 115 QLAGLPSGSSFFIIEKHIEIRLIKTEALIKILKALADY-----AVQGV--DEI 165
 QY 196 AIAATTNAMIKEIAKSPF-----VELMDTWTFQSGINTDGAHPVNTNRIVOSGDIISL 249
 Db 166 AEGKTEA---EIVAKIEYEMKKKGVTAHSPDTWVLTKGNALPHGTFGETRIKKGDLVLF 222
 QY 250 NTFPMIFGYTALERTLFCDHVDASLDIWEKNVAVHRLGLIKRGARCKDAIENEM 309
 Db 223 DLGAVHKGYCSDITRIVAFSDITDEOKIYDYLEAOVAADVKKYKIKASEIDLTARNI 282
 QY 310 YREMDLTKRSPFGYCHYGRAGVELREDIDTELKPGWVSMPEMVLPEGM 369
 Db 283 IRRAGGDFPRLHGLGLASVHEF-----PSITETNMELQENMVFTEPGIY----VP 333
 QY 370 GAGGYREHDLIVGEDAENITGFP 394
 Db 334 GVAGVRIEDDLVYTKDGVQLTEFP 358

RESULT 12

F65012
 hypothetical protein b2385 - *Escherichia coli* (strain K-12)
 C/Species: *Escherichia coli*
 C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
 C/Accession: F65012
 R/Batliner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; A.; Rose, D.J.; Mau, B.; Shao, Y.; Science 277, 1453-1462, 1997
 A/Title: The complete genome sequence of *Escherichia coli* K-12.
 A/Reference number: A64720; MUID:97426617; PMID:9278503
 A/Accession: F65012
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-361 <BLAT>
 A/Cross-references: UNIPROT:P76524; GB:AE000326; GB:U00096; NID:g1788718; PIDN:AACT544
 A/Experimental source: strain K-12, substrain MG1655
 C/Superfamily: X-Pro aminopeptidase

Query Match 10.5%; Score 229; DB 1; Length 361;
 Best Local Similarity 23.9%; Pred. No. 1.4e-11;

Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: C84047
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-364 <STO>
A:Cross-references: UNIPROT:Q9K828; GB:AP001518; GB:BA000004; NID:G10175792; PIDN:BA8066
A:Experimental source: strain C-125
C:Genetics:
A:Gene: pepQ
C:Superfamily: X-Pro aminopeptidase

Query Match 10.4%; Score 227.5; DB 2; Length 364;
Best Local Similarity 23.6%; Pred. No. 1.9e-11;
Matches 90; Conservative 65; Mismatches 171; Indels 55; Gaps 13;

```
QY 36 WMAKNNVDALFTSYHCINYSGLYC-----YGRKYGVNIDHNNATTI--SAG 83
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 10 WLKNDHSAFIQDKTSIFYLTF--YCDHERLVSLLPFAAPCLICPMETSIVKEAG 68
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 84 IDGGPWRSSFGDNITTYTMRDNPYRAVRLTGAARI---GIEFDHNLDFRRLQEE 139
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 69 -----WT---GEILGYSD--IEDPMLVROAVERKNAVLTSCTIVEAPRLTYARVALQD 117
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 140 ALPGVEFVDISQPSMMRTIKSLBEOKLIREGARVODVGAAACAAIKAGVPEHEVALAT 199
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 118 AFPALKLLIDGEPFLMELKQSAKELTTIKENALALDYGEVGVQHIQGRSEIETLALI 177
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 200 TNAMIREIAKSPFVFLMDTWTWFOGINTDGAHNPVTNRIVOSGDILSLNTPMI FGYY 259
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 178 EYELKRGKGVDMSEFGLTV-----LSGDQSANPHGNPGORTIKKGFVLFDLGVLDGYC 231
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 260 TALERTLCDDHYDASLDIWEKQNAVHRRGLELIKPGARCKDIAIELENTREMDLKKR 319
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 232 SDITRTVAFHVTDOODIYETVRKAQALDACRPGVEIRTIQIARTIITE----- 284
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 320 SFGYHSFGVLCHYYGRAGVELR-----DIDTE-LKPGWVVSMEPMVMLPEGMGAGG 373
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 285 -AGYGDYFP--HRIGHGLGMEVHELPSLNETNTDLQKGNVFTIEPGIYL---PSIGG 336
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 374 YREHDILIVGEDGAENITGFP 394
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 337 VRIEDDVVITBDGYQTLTNYP 357
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
```

Search completed: July 8, 2005, 01:36:11
Job time : 25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OW protein - protein search, using sw model

Run on: July 7, 2005, 23:34:26 ; Search time 84 Seconds

(without alignments)
2462.857 Million cell updates/sec

Title: US-10-807-228a-1

Perfect score: 2190

Sequence: 1 MTDDMLHVMKHNKGEKDYSP.....DGAENITGFPRFPHNIIIRN 404

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : UniProt 03:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2190	100.0	404	2	Q9RHU9
2	1712	78.2	402	2	Q7SIB5
3	1386.5	63.3	403	2	Q9R9T5
4	1376.5	62.9	403	2	Q9WVZ1
5	1350.5	61.7	403	2	Q88GQ1
6	1349.5	61.6	403	1	CREA_FIASU
7	1349	61.6	410	1	CREA_BACB0
8	1347.5	61.5	403	1	CREA_PSEBU
9	1329.5	60.7	411	2	O661B2
10	1279.5	12.8	376	2	Q8BML3
11	1274.5	12.5	364	2	O65KCI
12	1273.5	12.5	357	2	O67N93
13	1270	12.3	353	2	O65HH3
14	1261.5	11.9	365	2	Q72ZC0
15	1261.5	11.9	365	2	Q6HCR7
16	1258.5	11.8	365	2	O817E1
17	1257	11.7	363	2	O316B9
18	1255.5	11.7	353	2	O730Z5
19	1255.5	11.7	355	2	O74BM0
20	1255.5	11.7	365	2	O633T3
21	1255	11.6	353	1	YQHT_BACSU
22	1250.5	11.4	365	2	O81KX5
23	1248.5	11.3	351	2	O9UZP6
24	1248.5	11.3	353	2	O8EQ27
25	1247.5	11.3	353	2	O81M33
26	1247.5	11.3	353	2	O6HMD7
27	1247	11.3	355	2	O9K950
28	1247	11.3	351	1	O97BA0
29	1246.5	11.3	351	1	PEPO_PYRHO
30	1245.5	11.2	353	2	O634Y6
31	1243	11.1	360	2	Q9HJD2

32	241.5	11.0	353	2	O818P9	O818P9 bacillus ce
33	235.5	10.8	354	2	O836X1	O836X1 enterococcu
34	235.5	10.8	365	2	O71Z89	O71Z89 listeria mo
35	234	10.7	421	2	O67R80	O67R80 symbiodace
36	233.5	10.7	356	2	O636F0	O636F0 symbiodace
37	233	10.6	355	2	O8RAE3	O8RAE3 thermocaneer
38	232.5	10.6	356	2	O81WG2	O81WG2 bacillus an
39	232	10.6	406	2	O9KC35	O9KC35 bacillus ha
40	231.5	10.6	365	2	O92BD7	O92BD7 listeria in
41	230.5	10.5	365	2	O8Y6V3	O8Y6V3 listeria mo
42	230	10.5	356	2	O819T2	O819T2 bacillus ce
43	229.5	10.5	356	2	O6HET9	O6HET9 bacillus th
44	229	10.5	361	1	YPDF_ECOLI	YPDF_ECOLI escherichia
45	229	10.5	361	2	O8XBP9	O8XBP9 escherichia

ALIGNMENTS

RESULT 1					
Q9RHU9	PRELIMINARY;	PRT;	404 AA.		
AC	Q9RHU9;				
DT	01-MAY-2000 (TREMBLrel. 13, Created)				
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)				
DT	01-JUN-2003 (TREMBLrel. 24, Last annotation update)				
DE	Creatine amidinohydrolase.				
OS	Alcaligenes sp.				
OC	Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;				
OC	Alcaligenaceae; Alcaligenes.				
OX	NCBI_Taxid=512;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=KS-85;				
RA	Furukawa K., Ichikawa T., Koyama Y., Suzuki M.;				
RL	Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; AB016788; BAA88830.1; -				
DR	HSSP; P38488; ICHM.				
DR	GO; GO:0016980; F:creatinase activity; IEA.				
DR	GO; GO:0016787; F:hydrolase activity; IEA.				
DR	GO; GO:0008235; F:metalloexopeptidase activity; IEA.				
DR	GO; GO:0006500; P:creatinase metabolism; IEA.				
DR	GO; GO:0006508; P:proteolysis and peptidolysis; IEA.				
DR	InterPro; IPR000587; Creatinase.				
DR	InterPro; IPR000994; Peptidase_M24.				
DR	Pfam; PF01321; Creatinase_N; 1.				
DR	Pfam; PF00557; Peptidase_M24; 1.				
KW	Hydrolase.				
SC	SEQUENCE 404 AA; 46032 MW; 82A941891A1ABE79 CRC64;				
Query Match					
Best Local Similarity 100.0%; Score 2190; DB 2; Length 404;					
Matches 404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1	MTDDMLHVMKHNKGEKDYSPSPDAEMTRRONDVAGMAKNVDAALFTSYHCINYSGWL	60		
DB	1	MTDDMLHVMKHNKGEKDYSPSPDAEMTRRONDVAGMAKNVDAALFTSYHCINYSGWL	60		
QY	61	YCYFGRKYGVNIDHNNATTSAGIDGQPMRRSPGDNITTYDWRDNPYRAVROLTTGAK	120		
DB	61	YCYFGRKYGVNIDHNNATTSAGIDGQPMRRSPGDNITTYDWRDNPYRAVROLTTGAK	120		
QY	121	RIGIEPQVNLDRRQLEALPGYEFIDISQPSMMWRTISLEQKILREGARCDVGA	180		
DB	121	RIGIEPQVNLDRRQLEALPGYEFIDISQPSMMWRTISLEQKILREGARCDVGA	180		
QY	181	ACAAATKAGVBEHVAATATNAMIREFAKSPFVELADDTWFOGINTDGAHPVTR	240		
DB	181	ACAAATKAGVBEHVAATATNAMIREFAKSPFVELADDTWFOGINTDGAHPVTR	240		
QY	241	VQSGDILSLNTFPPIFGYATLERTLLFCDHVDDASLDIWEKNVAVHRGELILRGARCK	300		
DB	241	VQSGDILSLNTFPPIFGYATLERTLLFCDHVDDASLDIWEKNVAVHRGELILRGARCK	300		

QY 301 DIATELMEYREMDLTKRSPFGYSHSPGVLCYYYGREAGVELREDIDTELKPGWVNSMEP 360
 DB 301 DIATELMEYREMDLTKRSPFGYSHSPGVLCYYYGREAGVELREDIDTELKPGWVNSMEP 360
 QY 361 MVMLEPGMPGAGYREHDLIVGEDGAGENITGPPGPEHNIRN 404
 DB 361 MVMLEPGMPGAGYREHDLIVGEDGAGENITGPPGPEHNIRN 404

RESULT 2

Q7SIBS PRELIMINARY; PRT; 402 AA.
 ID 07SIBS
 AC 07SIBS
 DT 01-MAR-2004 (TrEMBLrel. 26, Created)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hydrolyase.
 OS Actinobacillus.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae.
 NC NCB1_TaxID=713;
 RN [1]
 RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).
 RX PubMed:12136144; DOI=10.1107/S0907444902010156;
 RA Padmanabhan B., Paehler A., Horikoshi M.;
 RL "Structure of creatine amidohydrolyase from Actinobacillus";
 RL Acta Crystallogr. D Biol. Crystallogr. 58:1322-1328 (2002).
 DR PDB; 1KPO; X-ray; A/B=1-402.
 DR GO; GO:0016980; F:creatinase activity; IEA.
 DR GO; GO:0008235; F:metalloxopeptidase activity; IEA.
 DR GO; GO:0006600; P:creatin metabolism; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR000587; Creatinase.
 DR InterPro; IPR000994; Peptidase_M24.
 DR Pfam; PF01321; Creatinase_N; 1.
 DR Pfam; PF00557; Peptidase_M24; 1.
 DR 3D-structure; Hydrolyase.
 SQ SEQUENCE 402 AA; 45872 MW; 2F6B95BA90938B5 CRC64;

Query Match 78.2%; Score 1712; DB 2; Length 402;

Best Local Similarity 75.2%; Pred. No. 1,2e-129;
 Matches 300; Conservative 50; Mismatches 49; Indels 0; Gaps 0;

QY 5 MLHMKNHNEKXSPSPDAEMTRONDVGMKANNVDALFTSYHCINYSGLYCYR 64
 DB 3 MIZTKZHNGZKKTTPSSZAEATRZBLRAMMAKSIDAVLFTSYHNINYSGLYCYR 62
 QY 65 GRKXGMYIDHNNATTISAGIDGQPMRRSPFGDNITTYDMRRDNFYRAVROLTTGAKRIGI 124
 DB 63 GRKXAZIYIBKAVTISKIGDGMRRSRGBOINVTYDMKRDNFTYSAVKLVGAKKIGI 122
 QY 125 EFDHVIDFRRQLEALPGVEFYVDISQSMWRTIKSLEROKLIREGARVCDVGAACAA 184
 DB 123 EHDHVTLBHRRZLZKALPGTEFYVDVGPVMWZRVYIKSSEZBLIRZGARISDIGAATAA 182
 QY 185 AIKAGVEHVAATTAAMTRELAKSPFVELMDTWTFOSGINTDGAHPVNRIVQSG 244
 DB 183 AIGSGVEYEVAAITABMARZARBPVYELMDTWTFOSGINTDGAHPVNRIVZRG 242
 QY 245 DILSLNFPPIFGYTYALERTLPCDHVDASLDIWEKNVAVHRRGELIRPGARCKDIAI 304
 DB 243 DILSLNFPPIFGYTYALERTLFLZZVBDASLZIWZKATAVHRGELIRPGARCKDIAI 302
 QY 305 ELNEMRYEMDLTKRSPFGYSHSPGVLCYYYGREAGVELREDIDTELKPGWVNSMEP 364
 DB 303 ELNEMRYEMDLTKRSPFGYSHSPGVLCYYYGREAGVELREDIDTELKPGWVNSMEP 362
 QY 365 PEGEPGAGYREHDLIVGEDGAGENITGPPGPEHNIR 403
 DB 363 PEGEPGAGYREHDLIVGEDGAGENITGPPGPEHNIR 401

RESULT 3

Q9R9T5 PRELIMINARY; PRT; 403 AA.
 ID Q9R9T5
 AC Q9R9T5
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Creatinase.
 GN Name-cln.
 OS Pseudomonas putida.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 NC NCB1_TaxID=303;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RS65;
 RA Liu W.-H., Tang T.-Y., Wen C.-J., Lin Y.-C.;
 RT "Cloning, Sequencing, and Expression of the Creatinase Gene from
 RT Pseudomonas putida RS65 in Escherichia coli.";
 RL Food Sci. Agric. Chem. 1:115-121 (1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RS65;
 RA Tang T.-Y., Liu W.-H.;
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF170566; AAD52565.4; -
 DR HSSP; P38488; 1CHM.
 DR GO; GO:0016980; F:creatinase activity; IEA.
 DR GO; GO:0016787; F:hydrolyase activity; IEA.
 DR GO; GO:0008235; F:metalloxopeptidase activity; IEA.
 DR GO; GO:0006600; P:creatin metabolism; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR000587; Creatinase.
 DR InterPro; IPR000834; Peptidase_M14.
 DR InterPro; IPR000994; Peptidase_M24.
 DR Pfam; PF01321; Creatinase_N; 1.
 DR Pfam; PF00557; Peptidase_M24; 1.
 DR POSITE; PS00133; CARBOXYPEPT_ZN_2; UNKNOWN_1.
 KW Hydrolyase.
 SQ SEQUENCE 403 AA; 45680 MW; D082072C692C1A9A CRC64;

Query Match 63.3%; Score 1386.5; DB 2; Length 403;

Best Local Similarity 66.1%; Pred. No. 2,2e-103;
 Matches 259; Conservative 48; Mismatches 84; Indels 1; Gaps 1;

QY 13 NGEKXSPSPDAEMTRONDVGMKANNVDALFTSYHCINYSGLYCYRPGKXGMYI 72
 DB 11 NGEKXSPSPDAEMTRONDVGMKANNVDALFTSYHCINYSGLYCYRPGKXGMYI 70
 QY 73 DHNNATTISAGIDGQPMRRSPFGDNITTYDMRRDNFYRAVROLTTGAKRIGIEFDHVN 131
 DB 71 TQNAVVISANIDGQPMRRSPFGDNITTYDMRRDNFYRAVROLTTGAKRIGIEFDHVN 130
 QY 132 DFRQLEALPGVEFYVDISQSMWRTIKSLEROKLIREGARVCDVGAACAAIAKAGV 191
 DB 131 MNRDKLASRYQALELVDIAAPCRMRMIXSAEHAHAIROGARVADIGAAVVALDQVP 190
 QY 192 EHEVALATTAMTRELAKSPFVELMDTWTFOSGINTDGAHPVNRIVQSGDIISLNT 251
 DB 191 EHEVALATTAMTRELAKSPFVELMDTWTFOSGINTDGAHPVNRIVQSGDIISLNT 250
 QY 252 FPMIFGYTYALERTLPCDHVDASLDIWEKNVAVHRRGELIRPGARCKDIAIENEMYR 311
 DB 251 FPMIFGYTYALERTLPCDHVDASLDIWEKNVAVHRRGELIRPGARCKDIAIENEMYR 310
 QY 312 EMDLTKRSPFGYSHSPGVLCYYYGREAGVELREDIDTELKPGWVNSMEP 371
 DB 311 EMDLTKRSPFGYSHSPGVLCYYYGREAGVELREDIDTELKPGWVNSMEP 370
 QY 372 GYREHDLIVGEDGAGENITGPPGPEHNIR 403
 DB 371 GYREHDLIVGEDGAGENITGPPGPEHNIR 402

```
RESULT 4
Q9WMZ1 PRELIMINARY: PRT: 403 AA.
AC Q9WMZ1.
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Creatinase (EC 3.5.3.3).
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NTU-8;
RX MEDLINE=99247136; PubMed=10230521; DOI=10.1023/A:1018705831622;
RA Hong M.C., Chang J.C., Wu M.L., Chang M.C.;
RT "Expression and export of Pseudomonas putida NTU-8 creatinase by
RT Escherichia coli using the chitinase signal sequence of Aeromonas
RT hydrophila.";
RL Biochem. Genet. 36:407-415(1998).
DR EMBL; AF072304; AAD37463.1; -.
DR HSSP; P38488; 1CM.
DR GO; GO:0016980; F:creatinase activity; IEA.
DR GO; GO:0016787; F:hydrolyase activity; IEA.
DR GO; GO:0008235; F:metalloexopeptidase activity; IEA.
DR GO; GO:0006600; P:creatinase metabolism; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000587; Creatinase.
DR InterPro; IPR000834; Peptidase_M24.
DR Pfam; PF01321; Creatinase_N; 1.
DR Pfam; PF00557; Peptidase_M24.
DR PROSITE; PS00133; CARBOXYPEPT_ZN_2; UNKOWN_1.
KW Hydrolyase.
SQ SEQUENCE 403 AA; 45691 MW; 0790A5634E7F0E2 CRC64;

Query Match 62.9%; Score 1376.5; DB 2; Length 403;
Best Local Similarity 65.8%; Pred. No. 1.4e-102;
Matches 258; Conservative 48; Mismatches 85; Indels 1; Gaps 1;

QY 13 NGEKDYSPSDAEMTRONDVRCGMANKNVDAALFTSYHCINYYSGMLYCYGKRYGVI 72
DB 11 NGEKDYSPSDAEMTRONDVRCGMANKNVDAALFTSYHCINYYSGMLYCYGKRYGVI 70
QY 73 DHNNATTISAGIDGGQPMWRSGF-DNITYTDMRDNFRAVROLTTGAKRIGIEFDHVL 131
DB 71 TGEAVSISANIDGGQPMWRSGF-DNITYTDMRDNFRAVROLTTGAKRIGIEFDHVL 130
QY 132 DFRROLEALPGVEFDISQPSMMRTIKSLBEQKLIBGARVCDVGAAACAAIKAGVP 191
DB 131 MNRKLSRSYPOAEIVDIAAPCMRMKISAEHAIIRQGARVADIGAAVVEALRDVP 190
QY 192 EHEVAIATNMIREIAKSPFVELMDTWTFOSGINTDGAHPVTNRIVYOSGLISLNT 251
DB 191 EYEVAAHTQAMVREIARTYPSSELDWTWTFOSGINTDGAHPVTNRIVYOSGLISLNC 250
QY 252 FPMIFGYTALERTLFCDHVDASLDIWEKNVAVHRGELIKPGARKDIAIENEMYR 311
DB 251 FPMIFGYTALERTLFCDHVDASLDIWEKNVAVHRGELIKPGARKDIAIENEMYR 310
QY 312 EMDLTKRSFGYGSFGVLCHYRGREAGVELREDIDTELKQGVVSMPPVMTLPGMGGA 371
DB 311 EMDLTKRSFGYGSFGVLCHYRGREAGVELREDIDTELKQGVVSMPPVMTLPGMGGA 370
QY 372 GGYREHDLIVIGDGAENITGPPGGEHNIR 403
DB 371 GGYREHDLIVIGDGAENITGPPGGEHNIR 402
```

RESULT 5
Q88GQ1

```
Q88GQ1 PRELIMINARY: PRT: 403 AA.
AC Q88GQ1.
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Creatinase.
GN OrderedLocustNames=PP3667;
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=22423060; PubMed=12534463;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L.M., Beanan M.J., Deboy R.T., Daugherty S.C., Kolonay J.F.,
RA Madupu R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,
RA Hance I., Chits Lee P., Holtzapfle E.K., Scanlan D., Tran K.,
RA Moazzar A., Ulteback T.R., Rizzo M., Lee K., Kosack D., Moestl D.,
RA Wedler H., Lauber J., Stjepandic D., Hobeisel J., Straetz M., Helm S.,
RA Kiewitz C., Eisen J.A., Timmis K.N., Duesterhoeft A., Tsemmler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440.";
RL Environ. Microbiol. 4:799-808(2002).
DR EMBL; AE016787; AAN69267.1; -.
DR HSSP; P38488; 1CM.
DR TIGR; PP3667; -.
DR GO; GO:0016980; F:creatinase activity; IEA.
DR GO; GO:0016787; F:hydrolyase activity; IEA.
DR GO; GO:0008235; F:metalloexopeptidase activity; IEA.
DR GO; GO:0006600; P:creatinase metabolism; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000587; Creatinase.
DR InterPro; IPR000994; Peptidase_M24.
DR Pfam; PF01321; Creatinase_N; 1.
DR Pfam; PF00557; Peptidase_M24.
KW Complete proteome; Hydrolyase.
SQ SEQUENCE 403 AA; 45633 MW; 7946470E04727B80 CRC64;

Query Match 61.7%; Score 1350.5; DB 2; Length 403;
Best Local Similarity 63.2%; Pred. No. 1.7e-100;
Matches 253; Conservative 50; Mismatches 96; Indels 1; Gaps 1;

QY 5 MLAVMKHNGEKDYSPSDAEMTRONDVRCGMANKNVDAALFTSYHCINYYSGMLYCYF 64
DB 3 MPKTKIRNGDKVQPTSAOEYARHARLRAVMAEODIEAAIFTSYHNVYISDFLYCSF 62
QY 65 GRKYGVI DHNNATTISAGIDGGQPMWRSGF-DNITYTDMRDNFRAVROLTTGAKRIG 123
DB 63 GRPYALVTQDKVVISANIDGGQPMWRSGF-DNITYTDMRDNFRAVROLTTGAKRIG 122
QY 124 IEPDHVNLDFRROLEALPGVEFDISQPSMMRTIKSLBEQKLIBGARVCDVGAAACA 183
DB 123 VEYDHVNLDFRROLEALPGVEFDISQPSMMRTIKSLBEQKLIBGARVCDVGAAACA 182
QY 184 AATAGVPEHEVAIATNMIREIAKSPFVELMDTWTFOSGINTDGAHPVTNRIVYOS 243
DB 183 EALREQVDEYEVAAHTQAMVREIARTYPSSELDWTWTFOSGINTDGAHPVTNRIVYOS 242
QY 244 GDIISLNTFPMIFGYTALERTLFCDHVDASLDIWEKNVAVHRGELIKPGARKDIAI 303
DB 243 GEIISLNTFPMIFGYTALERTLFCDHVDASLDIWEKNVAVHRGELIKPGARKDIAI 302
QY 304 IELNEMYREMDLTKRSFGYGSFGVLCHYRGREAGVELREDIDTELKQGVVSMPPV 363
DB 303 IELNEMYREMDLTKRSFGYGSFGVLCHYRGREAGVELREDIDTELKQGVVSMPPV 362
QY 364 LPBGPAGAGYREHDLIVIGDGAENITGPPGGEHNIR 403
DB 363 LPBGPAGAGYREHDLIVIGDGAENITGPPGGEHNIR 402
```

RESULT 6
CREA_FLASU STANDARD; PRT; 403 AA.
ID CREA_FLASU
AC P19213;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Creatinase (EC 3.5.3.3) (Creatine amidinohydrolase).
OS Flavobacterium sp. (strain U-188).
OC Bacteria; Bacteroidetes; Flavobacteriales;
OC Flavobacteriaceae; Flavobacterium.
NCBI_TaxId=242;
RN NCBI_TaxId=242;
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-15.
RX MEDLINE=91103958; PubMed=1368564;
RA Koyama Y., Kltao S., Yamamoto-Otake H., Suzuki M., Nakano E.;
RT "Cloning and expression of the creatinase gene from Flavobacterium sp.
RT U-188 in Escherichia coli";
RL Agric. Biol. Chem. 54:1453-1457 (1990).
RN [2]
RP SEQUENCE FROM N.A.
RA Suzuki K., Sagai H., Sugiyama M., Imamura S.;
RT "Molecular cloning and high expression of the Bacillus creatinase gene
RT in Escherichia coli";
RL J. Ferment. Bioeng. 76:77-81 (1993).
CC -1- CATALYTIC ACTIVITY: Creatine + H(2)O = sarcosine + urea.
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SIMILARITY: To other bacterial creatinases. Also to peptidases of
CC superfamily M24.
CC -1- CAUTION: Ref.1 sequence differs from that shown due to a
CC frameshift in position 336.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D00656; BAA3239.1; -;
DR EMBL: D14464; BAA03359.1; -;
DR EMBL: D14464; BAA03360.1; ALT_TERM.
DR PIR: JH0134; JH0134.
DR HSSP: P38488; 1CHM.
DR MEROPS: M24.UNB; -;
DR InterPro: IPR000587; Creatinase.
DR InterPro: IPR000994; Peptidase_M24.
DR Pfam: PF01321; Creatinase_N; 1.
DR Pfam: PF00557; Peptidase_M24; 1.
KW Direct protein sequencing; Hydrolase.
FT ACT SITE 232 232 By similarity.
FT ACT SITE 232 232 By similarity.
SQ SEQUENCE 403 AA; 45780 MM; 754BDCFA56D962C CRC64;
Query Match 61.6%; Score 1349.5; DB 1; Length 403;
Best Local Similarity 64.2%; Pred. No. 2.1e-100;
Matches 257; Conservative 46; Mismatches 96; Indels 1; Gaps 1;
OY 5 MLHWKMGKNGKXSPFSDAEMTRQNDVRCGMKANNVDALFTSYHCINYSGLVYCF 64
DB 3 MPTLRIRNGEKVSTFSAQRYANRACRAHAALENDAAVFTSYHNINYSDELVYCSF 62
OY 65 GRKGMVIDHNNATTISAGIDGGQPMRRSPG-DNITTYDWRDNEFYRAVRQLTTGAKRIG 123
DB 63 GRPVALVTTQDDVVISANIDGGQPMRRVGTNDIVTWDQGDNYFAVIAQALPRARRIG 122
OY 124 IEPHVNLDFFROLEALPGVEFYDISQPSMMWRITKSLBEOKLIREGARVCDVGGAACA 183
DB 123 IEHHLNINQNDKLAARYPDALVDVAAACRMEMIKSAEHEHMTRHGARVADIGGAATV 182
OY 184 AAIKAGVEHEVAIATTNAMIREIKSPFVELMOTMTWFGSGITDGAHPVTRIVOS 243

Db 183 EALRDQVPEYEVVALHATQAMVRAIAETFDNVLEMDTWTWFOSGINTDGAHPVTRKVNK 242
OY 244 GDITSLNTPPMIRGYTALRRTLFCDHVDDASLDIMKKNVAVHRGELIKRGARCDIA 303
DB 243 GDITSLNCPFPKGYTALRRTLFCDHCDSDHLEMMQANVEHAGKLIKPMRCSDA 302
OY 304 IELNEMVRMDLLKYSFGYSGSPGVACHYRGAGVLEDEDITELPKGMVSMPEPMVM 363
DB 303 KELNEIRLKNDLQYRTFGYSGHSPGTSHTYGRERAGLEDEDITVLEPGMVSMEMPM 362
OY 364 LPEGMPAGGYREHDLIVGEDGANTGPPGPEHNIIR 403
DB 363 LPEGRPGAGYREHDLIVNENGAENITKPPYGERNIIR 402
RESULT 7
CREA_BACBO STANDARD; PRT; 410 AA.
ID CREA_BACBO
AC P38487;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Creatinase (EC 3.5.3.3) (Creatine amidinohydrolase).
OS Bacillus sp. (strain B-0618).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
NCBI_TaxId=69000;
RN NCBI_TaxId=69000;
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-20.
RA Suzuki K., Sagai H., Sugiyama M., Imamura S.;
RT "Molecular cloning and high expression of the Bacillus creatinase gene
RT in Escherichia coli";
RL J. Ferment. Bioeng. 76:77-81 (1993).
CC -1- CATALYTIC ACTIVITY: Creatine + H(2)O = sarcosine + urea.
CC -1- SUBUNIT: Homodimer.
CC -1- INDUCTION: By choline chloride.
CC -1- SIMILARITY: To other bacterial creatinases. Also to peptidases of
CC superfamily M24.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D14463; BAA03358.1; -;
DR PIR: I39809; I39809.
DR HSSP: P38488; 1CHM.
DR InterPro: IPR000587; Creatinase.
DR InterPro: IPR000994; Peptidase_M24.
DR Pfam: PF01321; Creatinase_N; 1.
DR Pfam: PF00557; Peptidase_M24; 1.
KW Direct protein sequencing; Hydrolase.
FT INIT MET 0
FT ACT SITE 239 239 By similarity.
FT ACT SITE 239 239 By similarity.
SQ SEQUENCE 410 AA; 46618 MM; AAA2BFC1067B1A22 CRC64;
Query Match 61.6%; Score 1349; DB 1; Length 410;
Best Local Similarity 63.9%; Pred. No. 2.4e-100;
Matches 253; Conservative 49; Mismatches 90; Indels 4; Gaps 1;
OY 12 HNGKXQVSPFSDAEMTRQNDVRCGMKANNVDALFTSYHCINYSGLVYCFPKKGMV 71
DB 14 NGGKXVPTFSKEMETRNRLREYMAKAGIDAVFTSYHNINYSDELVYSRSTALV 73
OY 72 IDHNNATTISAGIDGGQPMRRSPG-DNITTYDWRDNEFYRAVRQLTT-----GAKRIGIEFD 127
DB 74 VTQDKHTVTSANIDAGPMRRSPDENIVTDDMKRDNLVAVKAYLNGSPSSGLGVEND 133
OY 128 HVNLDFFROLEALPGVEFYDISQPSMMWRITKSLBEOKLIREGARVCDVGGAACAAIK 187
DB 134 HMTLDLRQVQDAPNTELVDSQAVVMGHMFKSDDEEIDILKNGARVADIGGAAVVEAIR 193

QY 188 AGVPEHEVALATNAMI REIAKSPFVFLMDTWTFOSGINTDGAHNPVTNRIVOSGDI 247
 DB 194 EGVPEHEVALHGEAMVREIARTYPHAELRDTIMVQSGINTDGAHNPVTNRIVOSGDI 253
 QY 248 SLNTPFMIFGYTALERTLPCDHVDASLDIMKKNVAHNRGLLEIKPGARCKDIAELN 307
 DB 254 SLNCFPMIAGYATLERTLFLBEVSDRHELMWCKVHRGLLEIKPGARCKDIAELN 313
 QY 308 EMVREMDLKYRSFGYSGFVLCCHYRGAGVLEEDIDTELKPGMVTSMEMVNLPG 367
 DB 314 EIVREMDLANRTFGYSGFVLSHYRGAGVLEEDIDTELKPGMVTSMEMVNLPG 373
 QY 368 MPGAGYRHDILIVGEDGAENITGPPFPEHNIIR 403
 DB 374 EPGAGYRHDILIVSENGENTITKPPFPEHNIIR 409

RESULT 8
 CREA_PSEPU STANDARD; PRT: 403 AA.

AC P38488;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Creatinase (EC 3.5.3.3) (Creatine amidinohydrolyase).
 OS Pseudomonas putida.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=303;
 RN [1]
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
 RX MEDLINE=90339496; PubMed=1696320;
 RA Hoeftken H.W., Knof S.H., Bartlett P.A., Huber R., Moellerling H.,
 RT "Enzymatic mechanism of creatine amidinohydrolyase as deduced from
 crystal structures.";
 RN J. Mol. Biol. 214:597-610 (1990).
 RL [2]
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
 RX MEDLINE=89125596; PubMed=3221393;
 RA Hoeftken H.W., Knof S.H., Bartlett P.A., Huber R., Moellerling H.,
 RT "Crystal structure determination, refinement and molecular model of
 creatine amidinohydrolyase from Pseudomonas putida.";
 RL J. Mol. Biol. 204:417-433 (1988).
 CC -1- CATALYTIC ACTIVITY: Creatine + H(2)O = sarcosine + urea.
 CC -1- SUBUNIT: Homodimer.
 CC -1- DOMAIN: Each monomer has two clearly defined domains. The small N-
 terminal domain (AA 1-161) and the large domain (AA 162-403). Each
 of the two active sites is made by residues of the large domain of
 one monomer and some residues of the small domain of the other
 monomer.
 CC -1- SIMILARITY: To other bacterial creatinases. Also to peptidases of
 superfamily M24.
 CC PDB: 1CHM; X-ray; A/B=2-402.
 DR InterPro; IPR000587; Creatinase.
 DR InterPro; IPR000994; Peptidase_M24.
 DR Pfam; PF01321; Creatinase_N; 1.
 DR Pfam; PF00557; Peptidase_M24; 1.
 KW 3D-structure; Hydrolase.
 FT ACT SITE 232 232
 FT STRAND 6 7
 FT HELIX 21 37
 FT TURN 38 39
 FT STRAND 42 45
 FT HELIX 48 55
 FT TURN 62 63
 FT STRAND 67 70
 FT STRAND 75 80
 FT HELIX 81 83
 FT TURN 84 85
 FT HELIX 86 90

FT STRAND 95 99
 FT TURN 101 102
 FT TURN 104 105
 FT HELIX 106 114
 FT STRAND 120 123
 FT TURN 125 127
 FT STRAND 129 129
 FT HELIX 130 139
 FT TURN 141 142
 FT STRAND 144 147
 FT HELIX 149 156
 FT TURN 157 157
 FT HELIX 161 184
 FT TURN 187 188
 FT STRAND 190 190
 FT STRAND 191 209
 FT HELIX 215 215
 FT STRAND 219 224
 FT HELIX 225 229
 FT TURN 231 232
 FT STRAND 235 236
 FT STRAND 239 239
 FT TURN 242 243
 FT STRAND 245 251
 FT STRAND 253 254
 FT TURN 255 256
 FT STRAND 257 258
 FT STRAND 261 267
 FT HELIX 272 291
 FT TURN 294 295
 FT STRAND 297 297
 FT HELIX 298 312
 FT TURN 313 313
 FT HELIX 315 317
 FT STRAND 318 318
 FT STRAND 324 325
 FT STRAND 328 330
 FT TURN 331 332
 FT STRAND 333 334
 FT TURN 337 338
 FT STRAND 340 340
 FT TURN 343 344
 FT STRAND 348 348
 FT TURN 351 352
 FT STRAND 354 357
 FT TURN 360 363
 FT STRAND 365 366
 FT TURN 368 369
 FT STRAND 371 374
 FT STRAND 376 382
 FT TURN 383 384
 FT STRAND 385 388
 FT HELIX 396 399
 FT STRAND 400 401
 SQ SEQUENCE 403 AA; 45536 MM; 21C2CD80BD017E0 CRC64;

Query Match 61.5%; Score 1347.5; DB 1; Length 403;
 Best Local Similarity 64.0%; Pred. No. 3e-100;
 Matches 256; Conservative 47; Mismatches 96; Indels 1; Gaps 1;

QY 5 MLHWKMGKNGEKDPSFSDAEMTRONDVRGMKANVDALFTSYHCINYSGLWLYCF 64
 DB 3 MKTIRINGDGVNSTFSAQEVANKQARLRAHLAENIDALFTSYHNINYSDFLYCSF 62
 QY 65 GRKYGMVIDHNNATTISAGIDGQFWRSSFG-DNITTYDWRDNFRAVRQLTTGAKRIG 123
 DB 63 GRPYALVTBEDVVISANIDGGQFWRRTVGTNDIVYTDWRDNYFAALQQALPRARRIG 122
 QY 124 IEPDVNIDPFRQLEALPGVEFVDISQPSMMWRTIKLEQKLRBQAVCDVGAACA 183
 DB 123 IEHDLNLQNRDKLAAKYDPAELVDVAACRMRMVTKSAEBHVMIRHGARRADIGAAV 182
 QY 184 AAKAGVEHEVALATNAMI REIAKSPFVFLMDTWTFOSGINTDGAHNPVTNRIVOS 243

```

DB      183  EALGQVEYEVYALATQAVRALADTFEDVELMTWTWTFQSGINTDGAHPVTTTRKNK 242
QY      244  GDILSLNTPFPMIFGYTALERTLFCDDHYDASLDIWEKNVAVHRRGLBLIKPGARCKDIA 303
DB      243  GDILSLNTPFPMIAGYTTALERTLFLDHCSDHLRLQWVVEHAGLKLKPGARCKDIA 302
QY      304  IELNMTYREMDLKTFRSGYGHSPVULCHYTGREGVLEIREIDPELKGWVNSPEPVM 363
DB      303  RELNMTFLPKHVDLQYRTFGYGHSPFTLSHYGREGVLEIREIDPELKGWVNSPEPVM 362
QY      364  LPEGMGAGYREHDLIVGEDGAEINITGFPFGEHNIIIR 403
DB      363  LPBGLPGAGYREHDLIVNENGAENITKFPYGEHNIIIR 402

```

RESULT 9

```

O66182  PRELIMINARY; PRT; 411 AA.
ID  O66182
AC  O66182;
DT  01-AUG-1998 (TREMBlrel. 07, Created)
DT  01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT  01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE  Creatinase.
OC  Arthrobacter sp.
OC  Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC  Micrococciaceae; Micrococcaceae; Arthrobacter.
OX  NCBI_TaxID=1667;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=TE1826;
RA  Nishiyama Y., Imanaka T.;
RT  "Analysis of a negative regulator, SoxR, for the Arthrobacter
RL  J. Ferment. Bioceng. 81:64-67(1996).
RN  [2]
RP  SEQUENCE FROM N.A.
RC  STRAIN=TE1826;
RA  MEDLINE=96223334; PubMed=9563845; DOI=10.1007/s004380050685;
RX  Nishiyama Y., Toda A., Imanaka T.;
RT  "Gene cluster for creatinine degradation in Arthrobacter sp. TE1826.";
RL  Mol. Genet. 257:581-586(1998).
RN  [3]
RP  SEQUENCE FROM N.A.
RC  STRAIN=TE1826;
RA  Nishiyama Y., Imanaka T.;
RT  "Cloning and sequencing of the sarcosine oxidase gene from
RL  Arthrobacter sp. TE1826.";
RL  J. Ferment. Bioceng. 75:239-244(1993).
DR  BMBL; AB007122; BAA25928.1; -.
DR  HSSP; P38488; ICHM.
DR  GO; GO:0016980; F:creatinase activity; IEA.
DR  GO; GO:0016787; F:hydrolase activity; IEA.
DR  GO; GO:0008235; F:metalloexopeptidase activity; IEA.
DR  GO; GO:0006600; P:creatinase metabolism; IEA.
DR  GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR  InterPro; IPR000587; Creatinase.
DR  InterPro; IPR000994; Peptidase M24.
DR  Pfam; PF01321; Creatinase N; 1.
DR  Pfam; PF00557; Peptidase M24; 1.
KW  Hydrolase.
SQ  SEQUENCE 411 AA; 47147 MW; B8C5DS9A886DE201 CRC64;

```

Query Match 60.7%; Score 1329.5; DB 2; Length 411;
 Best Local Similarity 63.0%; Pred. No. 8.9e-99;
 Matches 250; Conservative 55; Mismatches 87; Indels 5; Gaps 3;

```

QY      12  HNGEADYSPFSAEMTRNDVGRGMKKNVDAALFTSYHCINYYSGMLYCFGRKGYGV 71
DB      14  HNGEKFQGFSGKEMDRNTNLRNVAEKDIDAVLFTSYHNINYYSDFLYTSFNNYGLV 73
QY      72  IDHNAATTSAGIDGQPFRRSFGDNITVTWRDDFYRAVROL--TTGAK--RIGIEDP 127

```

```

DB      74  VTQNKVTVSANIDGKPMWRSSYDENITYTDMRDNYFVALQKYLEAGVKKARLIGIED 133
QY      128  HVNIDFRQLEBALPGVEFVDISQPSMMNRITKLEBCKLIREGARVCDVGGAACAAAIK 187
DB      134  HVSIDLKRFSDTFEPNELVAVSODVMKQRIKAAELRHKNGARLINDIGYAVVAIQ 193
QY      188  AGVEPEEVALATYATNAMIKEIKSPFVELMDTWTWTFQSGINTDGAHPVTTTRKNK 247
DB      194  EGVEPEEVALAGSQTMRREIKLYPQSELRRTWTFQSGINTDGAHSAWTSKQVKEIL 253
QY      248  SLNTPFPMIFGYTALERTLFCDDHYDASLDIWEKNVAVHRRGLBLIKPGARCKDIAELN 307
DB      254  SLNTPFPMIAGYTTALERTLFLAEVSDAHLKYWEINVEYHKGGLBLIKPGAVCKICABLN 313
QY      308  EMYREMDLKTFRSGYGHSPVULCHYTGREGVLEIREIDPELKGWVNSPEPVM 367
DB      314  EMPREHDLVKNRTFGYGHSPFTLSHYGREGVLEIREIDPELKGWVNSPEPVM 373
QY      368  MPMGAGYREHDLIVGEDGAEINITGFPFGEHNIIIR 403
DB      374  QPMGAGYREHDLIVQENGVVEDITGFPFGEHNIIIR 410

```

RESULT 10

```

O66183  PRELIMINARY; PRT; 376 AA.
ID  O66183
AC  O66183;
DT  01-MAR-2003 (TREMBlrel. 23, Created)
DT  01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT  01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE  Cobalt dependent X-Pro dipeptidase (EC 3.4.13.9).
DN  OrderedLocusNames=OB2829;
OS  Oceanobacillus theysensis.
OC  Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.
OX  NCBI_TaxID=182710;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=HTE831;
RA  MEDLINE=22220767; PubMed=12253576; DOI=10.1093/nar/gkf526;
RX  Takami H., Takaki Y., Uchiyama I.;
RT  "Genome sequence of Oceanobacillus theysensis isolated from the Iheya
RL  Ridge and its unexpected adaptive capabilities to extreme
RT  environments.";
RL  Nucleic Acids Res. 30:3927-3935(2002).
DR  BMBL; AP004602; BAC14785.1; -.
DR  GO; GO:0004251; F:X-Pro dipeptidase activity; IEA.
DR  GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR  InterPro; IPR000994; Peptidase M24.
DR  Pfam; PF00557; Peptidase M24; 1.
KW  Complete proteome; Hydrolase.
SQ  SEQUENCE 376 AA; 42679 MW; 046EAB64D13E0AD2 CRC64;

```

Query Match 12.8%; Score 279.5; DB 2; Length 376;
 Best Local Similarity 23.4%; Pred. No. 4.3e-14;
 Matches 92; Conservative 74; Mismatches 183; Indels 45; Gaps 10;

```

QY      29  RQNDVRGMKKNVDAALFTSYHCINYYSGMLYCFGRKGYGVIDHNAATTSAGIDGQ 88
DB      6  RVDNRLALMEBSNIDLTIMNVENQYINGKATYSPYVALDSKSLTIPSEENH 65
QY      89  PWRSS-----FGDNITYTDMRDNYFRAVROLTTGAKRIGIEDHVNLDNR 134
DB      66  AKETIDANELYIYHETNLYGDKSYDY----LETVISQYPNGT-RVGVEFSSLSKMA 120
QY      135  RQLEBALPGVEFVDISQPSMMNRITKLEBCKLIREGARVCDVGGAACAAAIKAGVPEHE 194
DB      121  NILKEA--GPELINIDHRIAEARFKONDEELEMVOEAKGLSLAKSLLENVAQPIETWE 178
QY      195  VALATNAMIKEIKSPFVELMDTWTWTFQSGINT-----DGAHPVTTTRIVQSGDIL 247
DB      179  LDRHGTQLLFEEF--DNYPDATL-DYFAPSPGGLERSIMPHYFSNTRKLLINDIVHSAQV 236
QY      248  SLNTPFPMIFGYTALERTLFCDDHYDASLDIWEKNVAVHRRGLBLIKPGARCKDIAELN 307

```


Db 227 ALN-----GYRACERTFFVGKPTDKDAPFAAAYEAQVAMKTAIVGTAKEDVAR 290
Qy 308 EHYREMDLTKRSFGYGSFGVLCHYGRBAGVELREDIDTLKSGMVSMPPMMLPBG 367
Db 291 NIFQGYNLEKCIHRTGHGIGI-----GHAEPSLRFDNDLVLOEGMTFTIEPGIYI--- 342
Qy 368 MPAGAGYREHDLIVGEDGAENITGPPGPEHNI 401
Db 343 -PGVGGFRHSDTVVLTNKGTKLVTEYPRIEHLI 375

RESULT 11
06SKC1 PRELIMINARY; PRT; 364 AA.
AC 06SKC1; 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Ykwy (Peptidase M24).
GN Name=Ykwy; ORFNames=BL03561, BL01594,
OS Bacillus licheniformis DSM 13.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=279010;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 13.
RX PubMed15383718;
RA Velth B., Herzberg C., Steckel S., Feesche J., Maurer K.H.,
RA Ehrenreich P., Baumeier S., Henne A., Liesegang H., Merkl R.,
RA Ehrenreich A., Gottschalk G.;
RT "The Complete Genome Sequence of Bacillus licheniformis DSM13, an
RT Organism with Great Industrial Potential.",
RL J. Mol. Microbiol. Biotechnol. 7:204-211(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 14580;
RA Rev M.W., Ranaiva P., Nelson B.A., Brody-Karpin S.D., Zaretsky E.J.,
RA Tang M., de Leon A.L., Xiang H., Guest V., Clausen I.G., Olsen P.B.,
RA Raasmussen M.D., Andersen U.T., Jorgensen P.L., Larsen T.S.,
RA Sorokin A., Bolotin A., Lapidus A., Galleron N., Ehrlich S.D.,
RA Berka R.M.;
RT "Complete genome sequence of the industrial bacterium Bacillus
RT licheniformis and comparisons with closely related Bacillus species.",
RL Genome Biol. 5:R77-R77(2004).
DR EMBL; AB017333; AA040493.1; -
DR EMBL; CP000002; AA023135.1; -
SQ SEQUENCE 364 AA; 40396 MW; 2ADD117D8D34B1C7 CRC64;

Query Match 12.5%; Score 274.5; DB 2; Length 364;
Best Local Similarity 23.3%; Pred. No. 11e-13;
Matches 94; Conservative 63; Mismatches 161; Indels 85; Gaps 11;

Qy 29 RQNDVRGMMAKNVDAALFTS-----YHCINYY----- 56
Db 3 RIEQISSLKKKNTSSAFHSEKENVFYLTGTEPHERLMGVITFQDEEPFVCQMGAG 62
Qy 57 ----SGWLICYFGKRYGVNIDHNNATTSAGIDGQPPRRSFGDNITYTWRDNFYRAV 112
Db 63 QARAGAGYFEIIG--YG---DHQN-----PM-----ELISSAL 90
Qy 113 RQLTTGARRIGIEPDHVNLDPRRQLEALPGVEPVDISQPSMMRTIKSLSEQKLIRESGA 172
Db 91 QKRNGQCLKVAVEKTLASRAEMLAQAVTNGAEIVASBERLNQIRVITDERKEISILREAA 150
Qy 173 RVCDVGAACAATKAGVPEHEVALATNNAMIR--EIAKSPFVELMDTWTWFOGINTDG 231
Db 151 KLAAYGVGAAGAAALKEGIAIDV-----AKIEYALKKK--GVQMSSTVTLRBEKSGQ 203
Qy 232 AHNPTVNRIVOSGDILSINTEPMIFGYTALERTLFCDHVDASLIDIKKNVAVHRRGLE 291
Db 204 PHGNPGRSTLRKGFVFLFDLGVITIDGYCSDIRTLIVYQNSBEKGEIYNTVTLQATEALK 263

Qy 292 LIKPGARCKDIAIEINEMTYREMDLTKRSFGYGSFGVLCHYGRBAGVELREDIDTLK 351
Db 264 MSKPGVRLGDDLDKARGLIEKAGYGDYFPNHLGHGLGSPHEX-----PSMHNDELK 318
Qy 352 PGVVSMPEPMVLPFGMGAGYREHDLIVGEDGAENITGPP 394
Db 319 QGMVYTIPEGIYVPE-----IGVRIEDDVLVTADGAAALTQVP 357

RESULT 12
067N93 PRELIMINARY; PRT; 357 AA.
AC 067N93; 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Xaa-Pro dipeptidase.
GN ORFNames=STH1865;
OS Symbiodacterium thermophilum.
OC Bacteria; Actinobacteria; Symbiodacterium.
OX NCBI_TaxID=2734;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IAM14863;
RA Ueda K., Yamashita A., Ishikawa J., Shimada M., Watsuji T.,
RA Morimura K., Ikeda H., Hattori M., Beppu T.;
RT "Complete genome sequence of an uncultured bacterium Symbiodacterium
RT thermophilum.",
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP006840; BAD04850.1; -
DR InterPro; IPR000994; Peptidase M24.
DR InterPro; IPR001714; Pept. M24 MAP.
DR Pfam; PF00557; Peptidase M24; 1.
DR PRINTS; PR00599; MAPPEPTIDASE.
KW Hydrolase.
SQ SEQUENCE 357 AA; 39193 MW; 43A1576729876460 CRC64;

Query Match 12.5%; Score 273.5; DB 2; Length 357;
Best Local Similarity 25.1%; Pred. No. 1.2e-13;
Matches 97; Conservative 65; Mismatches 175; Indels 49; Gaps 10;

Qy 22 SDAEMTRQNDVRGMMAKNVDAALFTSYHCINYYSGM-----LYCYFGKRYGVNIDHNN 76
Db 2 SQARLAR-----LRARLERGLDAVLAKPHRAVLISGFTSAGLLITADAAALVTDRY 57
Qy 77 ATTISAGIDGQPPRRSFGDNITYT---DWRDNFYRAVRLTTGARRIGIEPDHVNLD 132
Db 58 VEQAALAAQAPREPEVIAKESANQAVLTRIVEW-----GVKRIEGEGDFLTVD 103
Qy 133 PRQLEALPGVEPVDISQPSMMRTIKSLSEQKLIRESGAVCDVCGAACAAATKAGVPE 192
Db 104 EHGYQQPLGACQWTSVSGLVEELRMKIDETALMRRAAATBAEAFQIILPKGVIE 163
Qy 193 HEVAIATTNAMIRIASFPEVELMDTWTWFOGINTDGAHPVYTRIVOSGDILSLNTF 252
Db 164 RDLATELEYRHKKLGAGVAFETIV-----ASGARSLPHGVASDKALIEVGDLITPFG 217
Qy 253 PMIFGYTALERTLFCDHVDASLIDIKKNVAVHRRGLEIKPGARCKDIAIEINEMTRE 312
Db 218 AVYQGYCSDMTRFTWLTGEPDKREIYGIIVLEAKRGVAARPGITGR---ELDDVCRS 273
Qy 313 MDLKL-YRSF---GYGHSFGVLCHYGRBAGVELREDIDTLKPGMVSMPEPMVLPFGM 368
Db 274 YIAKGYREYFGHGTGEGVGYHIEGPRVS---QRGDVLVLRGMVTVTEPGIYL----- 325
Qy 369 PGAGYREHDLIVGEDGAENITGPP 394
Db 326 PGMGVRIEDMLVTEBSGASFTHSP 351

RESULT 13
065HH3 PRELIMINARY; PRT; 353 AA.
ID 065HH3

AC 065H3: 25-OCT-2004 (TReMBLrel. 28, Created)
DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE Yght (Peptidase M24B, X-Pro dipeptidase).
GN Name=yght; ORFNames=BL01543, BL102617;
OS Bacillus licheniformis DSM 13
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=279010;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 13;
RX PubMed=1583718;
RA Velth B., Herzberg C., Steckel S., Feesche J., Maurer K.H.,
RA Ehrenreich P., Baumeister S., Henne A., Liesegang H., Weickl R.,
RA Ehrenreich A., Gottschalk G.,
RA "The Complete Genome Sequence of Bacillus licheniformis DSM13, an
RT Organism with Great Industrial Potential."
RL J. Mol. Microbiol. Biotechnol. 7:204-211(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 14580;
RA Rev M.W., Ramaiah P., Nelson B.A., Brody-Karpin S.D., Zaretsky E.J.,
RA Tang M., de Leon A.L., Xiang H., Guest V., Clausen I.G., Olsen P.B.,
RA Rasmussen M.D., Andersen U.T., Jorgensen P.L., Larsen T.S.,
RA Sorek A., Bolotin A., Lepidus A., Galleron N., Ehrlich S.D.,
RA Berka R.M.;
RA "Complete genome sequence of the industrial bacterium Bacillus
RT licheniformis and comparisons with closely related Bacillus species";
RL Genome Biol. 5:R77-R77(2004).
DR EMBL: AE017333; AAU41491.1;
DR EMBL: CP000002; AAU24132.1;
DR EMBL: CP000002; AAU24132.1;
SQ SEQUENCE 353 AA; 38339 MW; 2EC5DCED99160647 CRC64;
Query Match 12.3%; Score 270; DB 2; Length 353;
Best Local Similarity 25.8%; Pred. No. 2,3e-13;
Matches 99; Conservative 58; Mismatches 162; Indels 64; Gaps 11;
QY 31 NDVGMKANNVDAALFTSYHCINYSGLYCYFGKRYGMVIDHNNATTISAGIDGGOPW 90
DB 4 NKLRELFGGLGIDGLVTSGVNLRITG---FTGSSGLAVISDQAVFIT----- 50
QY 91 RRSRGDNTTYDWRNDPFR-----AVRLTT-----GAKRIGIEFPHVNLDFRRQ 136
DB 51 -----DRTYEQADQIKSFYIOHAGGIVOKTAEVIKEMGINKIGFEOBKRYTGTAA 104
QY 137 LEEALPGVEFYDISOPSMWRITKSLEROKLIREGARVCDVGAACAARAGVPEHEVA 196
DB 105 YKEQLGAALVYPVSGSVKRLIKSSSEIKILKEAKIADAFSHILVIRPGITEIEVA 164
QY 197 IATTNAMIETIA--KSPFVELMDTWTFQSGINTDGAHNPVTNRIVOSGILSLNTPFM 254
DB 165 NELFMYRSGADHSSFDMI-----VASGVSSLPHGVAAGVAKIEKGLVTLDFPAY 216
QY 255 IFGYTALERTLFCCHVDASLDIWEKVAVHRRLGLIKPGARKDAIENEMYREMD 314
DB 217 YKGYCSDTIRTVAAVEPDELKRIYQTVFEADIGWRSIKPGITCK---PADATRY- 271
QY 315 LLKTRSFYGHSPGLCHYRGREAGVELRED-----IDTELKPGMVVSMPEMVLPEGM 368
DB 272 ---ISSQYGVDFG---HSTGHGLGMEVHESPALARSBDQMLEKGMVTVTFEGIYI--- 321
QY 369 PGAGGYRHDILIVGEDGAENIT 391
DB 322 PGKGGVRIEDDILVTEGNEISLT 344
RESULT 14
Q72ZCO PRELIMINARY; PRT; 365 AA.
Q72ZCO;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Proline dipeptidase (EC 3.4.13.9).
GN Name=pepQ; OrderedlocusNames=BC34748;
OS Bacillus cereus (Strain ATCC 10987).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=222523;
RN [1]
RP SEQUENCE FROM N.A.
RC PubMed=14960714; DOI=10.1093/nar/gkh258;
RA Raeko D.A., Ravel J., Oestad O.A., Helgason E., Cer R.Z., Jiang L.,
RA Shores K.A., Fouts D.E., Tourasse N.J., Anguilo S.V., Kolonay J.F.,
RA Nelson W.C., Kojstoe A.-B., Frazer C.M., Read T.D.;
RA "The genome sequence of Bacillus cereus ATCC 10987 reveals metabolic
RT adaptations and a large plasmid related to Bacillus anthracis pXO1";
RL Nucleic Acids Res. 32:977-988(2004).
DR EMBL: AE017279; AAS43649.1;
DR HSSP; BCB4748;
DR TIGR; P15034; IAI6.
DR GO; GO:0004239; F:methionyl aminopeptidase activity; IEA.
DR GO; GO:0004251; F:X-Pro dipeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000994; Peptidase M24.
DR InterPro; IPR001131; Peptidase M24B.
DR InterPro; IPR001714; Pept M24 MAP.
DR Pfam; PF00557; Peptidase M24; 1.
DR PRINTS; PR00599; MAPPEPTIDASE.
DR PROSITE; PS00491; PROLINE_PEPITIDASE; 1.
KW Complete proteome; Hydrolase.
SQ SEQUENCE 365 AA; 40636 MW; 9EC4FBC713B9BED CRC64;
Query Match 11.9%; Score 261.5; DB 2; Length 365;
Best Local Similarity 25.1%; Pred. No. 1.2e-12;
Matches 95; Conservative 67; Mismatches 187; Indels 29; Gaps 12;
QY 26 MTRRQNDVRCGMANNVDAALFTSYHCINYSGLYCY-FGRKRYGM-VIDHNNATTISAG 83
DB 1 MNARLEMLQMLKEKNEVAALFTSTPVFMTN-ICHPEPRLTGMFVPEKEPILCPK 59
QY 84 IDGGPFRBPFQDN---TTDWRNDPFRVAVQLTTGARIGIEFPHVNLDFRRQLEA 140
DB 60 MEEQANNAQMAHEITIGFTIDRPWDITAKIKDRGINAANAVALKEKHLNVEREELTKL 119
QY 141 LPGEFYDISOPSMWRITKSLEROKLIREGARVCDVGAACAARAGVPEHEVAIATT 200
DB 120 FPNAAFTSABEKRELRLIDKEKLSLRBAKADVAVEGVNAIKEDSELEVL----- 175
QY 201 NAMIREIAKSPFVELMDTWTFQSGINTDGAHNPVTNRIVOSGILSLNTPFMITGY 259
DB 176 -AIEHEHLKT-EGIHKMSFPTMVLGANSALPHGIPGANK-MKRGDFVLPDLGVITIDGYC 232
QY 260 TALERTLFCCHVDASLDIWEKVAVHRRLGLIKPGARKDAIENEMYRE---WDL 316
DB 233 SDITRTVAFDISSEQTRINTVLAAGLOAVEACKPVTYTGALIDNAARSVIADGYGDF 292
QY 317 KYRSGFYGHSPGLCHYRGREAGVELREDIDTELKPGMVVSMPEMVLPEGMGAGYRE 376
DB 293 PNR---LGHGLGISVHY-----PDVKAQNESPLKEGMVTVTFEGIYI---VPVVGVI 340
QY 377 HDILIVGEDGAENITGFP 394
DB 341 EDDIYIRKDSSEILTKFP 358

RESULT 15
Q6HCR7 PRELIMINARY; PRT; 365 AA.
Q6HCR7;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Xaa-Pro dipeptidase (Proline dipeptidase) (EC 3.4.13.9).
GN Name=pepQ; OrderedlocusNames=BT9727_4344;
OS Bacillus thuringiensis (subsp. konkukian).

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=180856;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-97-27;
 RA Brettn T.S., Bruce D., Challacombe J.F., Gilna P., Han C., Hill K.,
 RA Hitchcock P., Jackson P., Kelm P., Longmire J., Lucas S., Okinaka R.,
 RA Richardson P., Rubin E., Rice H.;
 RT "Complete genome sequence of *Bacillus thuringiensis* 97-27.";
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL, AE017355; AAT63552.1;
 DR GO; GO:0004239; F:methionyl aminopeptidase activity; IEA.
 DR GO; GO:0004251; F:X-Pro dipeptidase activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR000994; Peptidase_M24.
 DR InterPro; IPR001131; Peptidase_M24B.
 DR InterPro; IPR001714; Pept_M24_MAP.
 DR Pfam; PF00557; Peptidase_M24; 1.
 DR PRINTS; PR00599; MAPEPTIDASE.
 DR PROSITE; PS00491; PROLINE PEPTIDASE; 1.
 KW Complete proteome; Hydrolase.
 SQ SEQUENCE 365 AA; 40668 MW; 36A6B2971F3E5C6 CRC64;

Query Match 11.9%; Score 261.5; DB 2; Length 365;
 Best Local Similarity 24.9%; Pred.No.1.2e-12;
 Matches 94; Conservative 65; Mismatches 191; Indels 27; Gaps 10;

QY 26 MTRRQNDYRGMAKNNVDAALFTSYHCINYSGLICYFGKRYGM-VIDHNNATTISAGI 84
 Db 1 MNAIRENLMQMLKENQVEAAFLSTPNVFYMTNFCPEHERLGMFVQEKEPILICPM 60
 QY 85 DGGQWRRSFGDNI---TYTDWRNDNFYRAVRLTGAKRIGIEPDHVNLDPRQLEAL 141
 Db 61 EEGQARNAGMAHEIIGFTDTRPMDIAKAIKRGINANAVAIKEMHNVREYELTKLF 120
 QY 142 PGVEFVDISQPSMMWRITKSLERQKLREGARVCDVGAAACAALIKAGVPEHEVALATN 201
 Db 121 PNAAF TSAEKVRELRLIKDEKELSLREAAQADYAVEGVNAIKENRSELEVL----- 175
 QY 202 AMIREIASPPVELMDTWTFQSGINTDGAHN-PVTRIVQSGDILSLNTPPMIFGYT 260
 Db 176 AIIHEHLEKT-KGIHMSQSDTWLAGANSALPHGIPGANK-MNRGDPVLFDLGVIIIDGYCS 233
 QY 261 ALERTLPCDHYDDASLDIMEKVVAVHRRGLLEIKPGARCKDIAIENMYRE---WDLX 317
 Db 234 DITRTVAAGEISEEQTRIYNVYLAGQLAVEACKPGVTLGALDNAASVIADAGYGFPP 293
 QY 318 YRSFGYHSPVLCYYGREAGVELREDIDTELKPGMVSMEPMVMLPEGMPGAGYREH 377
 Db 294 HR---LGHGLGISVNEY-----PDVKAGNESPLKEGWFTTEPGIY---VPNVGVRIE 341
 QY 378 DILVIGEDGAENITGFP 394
 Db 342 DDYIYTKDGSSEILTKFP 358

Search completed: July 8, 2005, 01:35:42
 Job time : 87 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 8, 2005, 01:18:01 ; Search time 27 Seconds
(without alignments)
1116.971 Million cell updates/sec

Title: US-10-807-228A-1

Perfect score: 2190

Sequence: 1 MTDMDLHVKNHNGEKDYSF.....DGAENITGPFPGPHNIIRN 404

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 513545 seqs, 7464964 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: Issued Patents AA:*
2: /cgn2_6/prodata/1/aa/5A_COMB.pep:*
3: /cgn2_6/prodata/1/aa/6A_COMB.pep:*
4: /cgn2_6/prodata/1/aa/6B_COMB.pep:*
5: /cgn2_6/prodata/1/aa/PCTUS_COMB.pep:*
6: /cgn2_6/prodata/1/aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Query Match	Length	DB ID	Description
1	2190	100.0	404	1	US-09-940-941-1
2	2190	100.0	404	2	US-08-947-726A-2
3	2190	100.0	404	3	US-08-799-897-1
4	2190	100.0	404	4	US-09-856-645-1
5	2190	98.1	404	4	US-09-869-280A-1
6	236.5	10.8	362	4	US-09-134-000C-4306
7	219	10.0	366	4	US-09-107-532A-4292
8	200.5	9.2	370	3	US-09-134-001C-3769
9	196.5	9.0	390	3	US-09-134-001C-3627
10	188.5	8.6	360	4	US-09-583-110-3285
11	188.5	8.6	361	4	US-09-107-433-3390
12	187.5	8.6	363	4	US-09-602-777A-94
13	183	8.4	353	4	US-09-583-110-4793
14	183	8.4	358	4	US-09-107-433-3157
15	181.5	8.3	329	4	US-09-861-451A-36
16	171	7.8	355	4	US-09-198-452A-871
17	171	7.8	364	4	US-09-438-185A-815
18	160.5	7.3	253	4	US-09-523-263B-10
19	147	6.7	318	4	US-09-107-532A-5738
20	143.5	6.6	289	4	US-09-797-906-4
21	143	6.5	379	4	US-09-602-777A-112
22	142.5	6.5	289	4	US-09-797-906-5
23	142	6.5	274	4	US-09-523-263B-11
24	140.5	6.4	248	4	US-09-523-263B-15
25	140.5	6.4	485	4	US-09-902-540-9878
26	140	6.4	441	4	US-09-543-681A-5018
27	138	6.3	520	4	US-09-724-623-124

28	135	6.2	416	4	US-09-252-991A-26713	Sequence 26713, A
29	134	6.1	394	4	US-09-523-263B-2	Sequence 2, Appl
30	133	6.1	387	4	US-09-523-263B-9	Sequence 9, Appl
31	133	6.1	492	4	US-09-489-039A-11838	Sequence 11838, A
32	128.5	5.9	260	4	US-09-902-540-12329	Sequence 12329, A
33	128	5.8	264	4	US-09-328-352-7576	Sequence 7576, Ap
34	127.5	5.8	265	4	US-09-543-681A-7847	Sequence 7847, Ap
35	127.5	5.8	441	4	US-09-492-709A-272	Sequence 272, App
36	124.5	5.7	250	4	US-09-523-263B-22	Sequence 22, Appl
37	123.5	5.6	252	4	US-10-138-701-26	Sequence 26, Appl
38	121.5	5.5	389	4	US-09-270-767-41600	Sequence 41600, A
39	117	5.3	305	4	US-09-523-263B-14	Sequence 14, Appl
40	116	5.3	1452	2	US-08-652-971-4	Sequence 4, Appl
41	116	5.3	1452	2	US-08-991-258A-4	Sequence 4, Appl
42	116	5.3	1452	2	US-08-769-399-4	Sequence 4, Appl
43	116	5.3	1452	3	US-08-991-953A-4	Sequence 4, Appl
44	115.5	5.3	285	4	US-09-252-991A-20742	Sequence 20742, A
45	114.5	5.2	268	4	US-09-523-263B-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-09-940-941-1
; Sequence 1, Application US/09940941
; Patent No. RE38687
; GENERAL INFORMATION:
; APPLICANT: Sogabe, Atsushi
; Hattori, Takashi
; Nishiya, Yoshiaki
; Kawamura, Yoshihisa
; TITLE OF INVENTION: NOVEL CREATINE AMIDINOHYDROLASE, PRODUCTION
; THEREOF AND USE THEREOF
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LEYDIG, VOIT & MAYER, LTD.
; STREET: Two Prudential Plaza, Suite 4900
; CITY: Chicago
; STATE: Illinois
; COUNTRY: US
; ZIP: 60601-6780
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/940,941
; FILING DATE: 28-Aug-2001
; CLASSIFICATION DATA:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/799,897
; FILING DATE: 13-FEB-1997
; APPLICATION NUMBER: JP 25435/1996
; FILING DATE: 13-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Robert F. Green
; REGISTRATION NUMBER: 27555
; REFERENCE/DOCKET NUMBER: 78064
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 616-5700
; TELEFAX: (312) 616-5700
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 404 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: <Unknown>
; DESCRIPTION: protein
; ORIGINAL SOURCE:
; ORGANISM: Alcaligenes faecalis

STRAIN: TE3581 (FERM P-14237)
FEATURE:
NAME/KEY: mat peptide
LOCATION: 1 to 404
OTHER INFORMATION: protein having creatine amidino-
hydrolase activity
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-940-941-1

Query Match 100.0%; Score 2190; DB 1; Length 404;
Best Local Similarity 100.0%; Pred. No. 1e-228;
Matches 404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTDDMLHWKMKNGEKDYSPFSDAETRRONDVGMANKNDVALFTSYHCINYYSGWL 60
DB 1 MTDDMLHWKMKNGEKDYSPFSDAETRRONDVGMANKNDVALFTSYHCINYYSGWL 60
QY 61 YCYFGKRYGMVLDHNNATTISAGIDGQPMRRSPGDNITYTDRRDNFYRAVQLTTGAK 120
DB 61 YCYFGKRYGMVLDHNNATTISAGIDGQPMRRSPGDNITYTDRRDNFYRAVQLTTGAK 120
QY 121 RIGIEFDHVNLDPRQLEALPGVEFVDISQSPMMRTIKSLBEQKLIREGARVCVGA 180
DB 121 RIGIEFDHVNLDPRQLEALPGVEFVDISQSPMMRTIKSLBEQKLIREGARVCVGA 180
QY 181 ACAAAIKAGVPEHEVAIATTNAMIRESFPFVELMDTWTFQSGINTDGAHPVTNR1 240
DB 181 ACAAAIKAGVPEHEVAIATTNAMIRESFPFVELMDTWTFQSGINTDGAHPVTNR1 240
QY 241 VOSGDLISINTPFMIFGYTTALERTLFCDHVDASLDIWEKNVAVHRRGLEIKPGARCK 300
DB 241 VOSGDLISINTPFMIFGYTTALERTLFCDHVDASLDIWEKNVAVHRRGLEIKPGARCK 300
QY 301 DIAIELNEMRWMDLKYRSFGYGSFGVLCHYYGREAGVELREDIDTELKPGMVSMER 360
DB 301 DIAIELNEMRWMDLKYRSFGYGSFGVLCHYYGREAGVELREDIDTELKPGMVSMER 360
QY 361 MWMLPGMPGAGYREHDILIVGEDGAENITGPPFGEHNIRN 404
DB 361 MWMLPGMPGAGYREHDILIVGEDGAENITGPPFGEHNIRN 404

RESULT 2

US-08-947-726A-2
Sequence 2, Application US/08947726A
Patent No. 5932466

GENERAL INFORMATION:

APPLICANT: Furukawa, Keisuke
APPLICANT: Ichikawa, Toshio
APPLICANT: Suzuki, Masaru
APPLICANT: Koyama, Yasuji
TITLE OF INVENTION: A NOVEL CREATIVE AMIDINOHYDROLASE GENE,
TITLE OF INVENTION: A NOVEL RECOMBINANT DNA, AND A PROCESS FOR PRODUCING
TITLE OF INVENTION: CREATIVE AMIDINOHYDROLASE
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDWARDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/947,726A
FILING DATE: 09-OCT-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/535,444

FILING DATE: 27-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 94/235737
FILING DATE: 29-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lawrence III, Stanton T.
REGISTRATION NUMBER: 25,736
REFERENCE/DOCKET NUMBER: 7005-120-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 404 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-947-726A-2

Query Match 100.0%; Score 2190; DB 2; Length 404;
Best Local Similarity 100.0%; Pred. No. 1e-228;
Matches 404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTDDMLHWKMKNGEKDYSPFSDAETRRONDVGMANKNDVALFTSYHCINYYSGWL 60
DB 1 MTDDMLHWKMKNGEKDYSPFSDAETRRONDVGMANKNDVALFTSYHCINYYSGWL 60
QY 61 YCYFGKRYGMVLDHNNATTISAGIDGQPMRRSPGDNITYTDRRDNFYRAVQLTTGAK 120
DB 61 YCYFGKRYGMVLDHNNATTISAGIDGQPMRRSPGDNITYTDRRDNFYRAVQLTTGAK 120
QY 121 RIGIEFDHVNLDPRQLEALPGVEFVDISQSPMMRTIKSLBEQKLIREGARVCVGA 180
DB 121 RIGIEFDHVNLDPRQLEALPGVEFVDISQSPMMRTIKSLBEQKLIREGARVCVGA 180
QY 181 ACAAAIKAGVPEHEVAIATTNAMIRESFPFVELMDTWTFQSGINTDGAHPVTNR1 240
DB 181 ACAAAIKAGVPEHEVAIATTNAMIRESFPFVELMDTWTFQSGINTDGAHPVTNR1 240
QY 241 VOSGDLISINTPFMIFGYTTALERTLFCDHVDASLDIWEKNVAVHRRGLEIKPGARCK 300
DB 241 VOSGDLISINTPFMIFGYTTALERTLFCDHVDASLDIWEKNVAVHRRGLEIKPGARCK 300
QY 301 DIAIELNEMRWMDLKYRSFGYGSFGVLCHYYGREAGVELREDIDTELKPGMVSMER 360
DB 301 DIAIELNEMRWMDLKYRSFGYGSFGVLCHYYGREAGVELREDIDTELKPGMVSMER 360
QY 361 MWMLPGMPGAGYREHDILIVGEDGAENITGPPFGEHNIRN 404
DB 361 MWMLPGMPGAGYREHDILIVGEDGAENITGPPFGEHNIRN 404

RESULT 3

US-08-799-897-1
Sequence 1, Application US/08799897
Patent No. 6080553

GENERAL INFORMATION:

APPLICANT: Sogabe, Atsushi
APPLICANT: Hattori, Takashi
APPLICANT: Nishiyama, Yoshiaki
APPLICANT: Kawamura, Yoshihisa
TITLE OF INVENTION: NOVEL CREATIVE AMIDINOHYDROLASE, PRODUCTION
TITLE OF INVENTION: THEREOF AND USE THEREOF
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: LEYDIG, VOIT & MAYER, LTD.
STREET: Two Prudential Plaza, Suite 4900
CITY: Chicago
STATE: Illinois
COUNTRY: US
ZIP: 60601-6780

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/799,897
; FILING DATE: 13-FEB-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 25435/1996
; FILING DATE: 13-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Robert F. Green
; REGISTRATION NUMBER: 27555
; REFERENCE/DOCKET NUMBER: 78064
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 616-5600
; TELEFAX: (312) 616-5700
; TELEX: 25-3533
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 404 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Alcaligenes faecalis
; STRAIN: TE3581 (FERM P-14237)
; FEATURE:
; NAME/KEY: mat peptide
; LOCATION: 1 to 404
; OTHER INFORMATION: protein having creatine amidino-
; OTHER INFORMATION: hydrolase activity
US-08-799-897-1

Query Match      100.0%; Score 2190; DB 3; Length 404;
Best Local Similarity 100.0%; Pred. No. 1e-228;
Matches 404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTDDMLHVMKMHNGEADYSPFSDAEMTRONDVGMMAKNVDAALFTSYHCINYYSGWL 60
DB 1 MTDDMLHVMKMHNGEADYSPFSDAEMTRONDVGMMAKNVDAALFTSYHCINYYSGWL 60
QY 61 YCYFGRKYGWVIDHNNATTISAGIDGQPWRRSFGDNITYTDMRRDNFYRAVQLTTGAK 120
DB 61 YCYFGRKYGWVIDHNNATTISAGIDGQPWRRSFGDNITYTDMRRDNFYRAVQLTTGAK 120
QY 61 YCYFGRKYGWVIDHNNATTISAGIDGQPWRRSFGDNITYTDMRRDNFYRAVQLTTGAK 120
DB 61 YCYFGRKYGWVIDHNNATTISAGIDGQPWRRSFGDNITYTDMRRDNFYRAVQLTTGAK 120
QY 121 RIGIEFDHVNLDPRRLQLEALPGVEFVDISQPSMMRTIKSLBEOKLIBGARVCVCGA 180
DB 121 RIGIEFDHVNLDPRRLQLEALPGVEFVDISQPSMMRTIKSLBEOKLIBGARVCVCGA 180
QY 121 RIGIEFDHVNLDPRRLQLEALPGVEFVDISQPSMMRTIKSLBEOKLIBGARVCVCGA 180
DB 121 RIGIEFDHVNLDPRRLQLEALPGVEFVDISQPSMMRTIKSLBEOKLIBGARVCVCGA 180
QY 181 ACAAAIKAGVPEHEVAIATTNAMIREIAKSPFVELMDTWTFQSGINTDGANPVYTNRI 240
DB 181 ACAAAIKAGVPEHEVAIATTNAMIREIAKSPFVELMDTWTFQSGINTDGANPVYTNRI 240
QY 181 ACAAAIKAGVPEHEVAIATTNAMIREIAKSPFVELMDTWTFQSGINTDGANPVYTNRI 240
DB 181 ACAAAIKAGVPEHEVAIATTNAMIREIAKSPFVELMDTWTFQSGINTDGANPVYTNRI 240
QY 241 VOSGDILSLNTPFMIFGYTTALERTLFCDHVDDASDIWEKNAVARRGLELIKPGARCK 300
DB 241 VOSGDILSLNTPFMIFGYTTALERTLFCDHVDDASDIWEKNAVARRGLELIKPGARCK 300
QY 301 DIAIELNEMRYREMDLLKYSFGYGSFGVLCHYYGAEAGVELREDIDTELKPGMVVSMEP 360
DB 301 DIAIELNEMRYREMDLLKYSFGYGSFGVLCHYYGAEAGVELREDIDTELKPGMVVSMEP 360
QY 301 DIAIELNEMRYREMDLLKYSFGYGSFGVLCHYYGAEAGVELREDIDTELKPGMVVSMEP 360
DB 301 DIAIELNEMRYREMDLLKYSFGYGSFGVLCHYYGAEAGVELREDIDTELKPGMVVSMEP 360
QY 361 MMLPEGMGAGGYRHHDLIIYGEDGAENITGPPFGPHNIIRN 404
DB 361 MMLPEGMGAGGYRHHDLIIYGEDGAENITGPPFGPHNIIRN 404
; RESULT 4
; US-09-856-645-1
; Sequence 1, Application US/09856645
```

```

; Patent No. 6639700
; GENERAL INFORMATION:
; APPLICANT: KIRKMAN CORPORATION
; TITLE OF INVENTION: CREATINE AMIDINOHYDROLASE AND PROCESSES FOR PRODUCING THE SAME
; FILE REFERENCE: PH-688-PCT
; CURRENT APPLICATION NUMBER: US/09/856,645
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: JP10-334252
; PRIOR FILING DATE: 1998-11-25
; NUMBER OF SEQ ID NOS: 1
; SEQ ID NO 1
; LENGTH: 404
; TYPE: PRT
; ORGANISM: Alcaligenes sp.
US-09-856-645-1

Query Match      100.0%; Score 2190; DB 4; Length 404;
Best Local Similarity 100.0%; Pred. No. 1e-228;
Matches 404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTDDMLHVMKMHNGEADYSPFSDAEMTRONDVGMMAKNVDAALFTSYHCINYYSGWL 60
DB 1 MTDDMLHVMKMHNGEADYSPFSDAEMTRONDVGMMAKNVDAALFTSYHCINYYSGWL 60
QY 61 YCYFGRKYGWVIDHNNATTISAGIDGQPWRRSFGDNITYTDMRRDNFYRAVQLTTGAK 120
DB 61 YCYFGRKYGWVIDHNNATTISAGIDGQPWRRSFGDNITYTDMRRDNFYRAVQLTTGAK 120
QY 61 YCYFGRKYGWVIDHNNATTISAGIDGQPWRRSFGDNITYTDMRRDNFYRAVQLTTGAK 120
DB 61 YCYFGRKYGWVIDHNNATTISAGIDGQPWRRSFGDNITYTDMRRDNFYRAVQLTTGAK 120
QY 121 RIGIEFDHVNLDPRRLQLEALPGVEFVDISQPSMMRTIKSLBEOKLIBGARVCVCGA 180
DB 121 RIGIEFDHVNLDPRRLQLEALPGVEFVDISQPSMMRTIKSLBEOKLIBGARVCVCGA 180
QY 121 RIGIEFDHVNLDPRRLQLEALPGVEFVDISQPSMMRTIKSLBEOKLIBGARVCVCGA 180
DB 121 RIGIEFDHVNLDPRRLQLEALPGVEFVDISQPSMMRTIKSLBEOKLIBGARVCVCGA 180
QY 181 ACAAAIKAGVPEHEVAIATTNAMIREIAKSPFVELMDTWTFQSGINTDGANPVYTNRI 240
DB 181 ACAAAIKAGVPEHEVAIATTNAMIREIAKSPFVELMDTWTFQSGINTDGANPVYTNRI 240
QY 181 ACAAAIKAGVPEHEVAIATTNAMIREIAKSPFVELMDTWTFQSGINTDGANPVYTNRI 240
DB 181 ACAAAIKAGVPEHEVAIATTNAMIREIAKSPFVELMDTWTFQSGINTDGANPVYTNRI 240
QY 241 VOSGDILSLNTPFMIFGYTTALERTLFCDHVDDASDIWEKNAVARRGLELIKPGARCK 300
DB 241 VOSGDILSLNTPFMIFGYTTALERTLFCDHVDDASDIWEKNAVARRGLELIKPGARCK 300
QY 301 DIAIELNEMRYREMDLLKYSFGYGSFGVLCHYYGAEAGVELREDIDTELKPGMVVSMEP 360
DB 301 DIAIELNEMRYREMDLLKYSFGYGSFGVLCHYYGAEAGVELREDIDTELKPGMVVSMEP 360
QY 301 DIAIELNEMRYREMDLLKYSFGYGSFGVLCHYYGAEAGVELREDIDTELKPGMVVSMEP 360
DB 301 DIAIELNEMRYREMDLLKYSFGYGSFGVLCHYYGAEAGVELREDIDTELKPGMVVSMEP 360
QY 361 MMLPEGMGAGGYRHHDLIIYGEDGAENITGPPFGPHNIIRN 404
DB 361 MMLPEGMGAGGYRHHDLIIYGEDGAENITGPPFGPHNIIRN 404
; RESULT 5
; US-09-869-280A-1
; Sequence 1, Application US/09869280A
; GENERAL INFORMATION:
; APPLICANT: Fuyukawa, Keisuke
; APPLICANT: Koyama, Yasuji
; APPLICANT: Suzuki, Masaru
; TITLE OF INVENTION: THERMOSTABLE CREATINE AMIDINOHYDROLASE AND PROCESS FOR
; FILE REFERENCE: 04853-0075-00000
; CURRENT APPLICATION NUMBER: US/09/869,280A
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: PCT/JP99/07424
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: JP-11-033359
; PRIOR FILING DATE: 1999-01-01
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 1
; LENGTH: 404
; TYPE: PRT
; ORGANISM: Alcaligenes sp.
US-09-869-280A-1
```

Query Match 98.1%; Score 2149; DB 4; Length 404;
 Best Local Similarity 98.0%; Pred. No. 2,9e-224;
 Matches 396; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

1 MTDDMLAHMKHNGKDYSPSDAEMTRNDVRGMMAKNVDAALFTSYHCINYSGL 60
 1 MTDDMLAHMKHNGKDYSPSDAEMTRNDVRGMMAKNVDAALFTSYHCINYSGL 60
 61 YCFGRKGMVNDHNNATTSAGIDGQPMWRSPGDNITTYTDMRDNFRAVROLTTGAK 120
 61 YCFGRKGMVNDHNNATTSAGIDGQPMWRSPGDNITTYTDMRDNFRAVROLTTGAK 120
 121 RIGIFDHYNDLFRQLEBALPGVEFVDISQPSMMWRITKSLEOKLIREGARVCDVGA 180
 121 RIGIFDHYNDLFRQLEBALPGVEFVDISQPSMMWRITKSLEOKLIREGARVCDVGA 180
 181 ACAAIKXGVPHEVAITNTMIREIAKSPFVELMDTWTFQSGINTDGAHNPVTNRI 240
 181 ACAAIKXGVPHEVAITNTMIREIAKSPFVELMDTWTFQSGINTDGAHNPVTNRI 240
 241 VQSGDILSLNTPMIFGYTTALERTLFCDHVDASLIDIEKNVAVHRGELIKPGARCK 300
 241 VQSGDILSLNTPMIFGYTTALERTLFCDHVDASLIDIEKNVAVHRGELIKPGARCK 300
 301 DIAIELNMYREMDLLKTRSPGYGHSFGVLYCHYGRAGVELREDIDTELKPGMYVMEP 360
 301 DIAIELNMYREMDLLKTRSPGYGHSFGVLYCHYGRAGVELREDIDTELKPGMYVMEP 360
 361 MVMLEPGMGAGYREHDLIYGEDEGAENITFPFGEENITRN 404
 361 MVMLEPGMGAGYREHDLIYGEDEGAENITFPFGEENITRN 404

RESULT 6
 US-09-134-000C-4306
 ; Sequence 4306, Application US/09134000C
 ; Patent No. 6617156

GENERAL INFORMATION:
 APPLICANT: Lynn Doucette-Stamm et al
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: 032/96-032
 CURRENT APPLICATION NUMBER: US/09/134,000C
 PRIOR FILING DATE: 1998-08-13
 PRIOR APPLICATION NUMBER: US 60/055,778
 FILING DATE: 1997-08-15
 NUMBER OF SEQ ID NOS: 6812
 SOFTWARE: Patentin version 3.1
 SEQ ID NO 4306
 LENGTH: 362
 TYPE: PRT
 ORGANISM: Enterococcus faecalis
 US-09-134-000C-4306

Query Match 10.8%; Score 236.5; DB 4; Length 362;
 Best Local Similarity 25.0%; Pred. No. 8.4e-17;
 Matches 99; Conservative 59; Mismatches 163; Indels 75; Gaps 13;

26 MTRRNDVRGMMAKNVDAALFTSYHCINYSGLYCYGRKGMVNDHNNATTSAGID 85
 9 MVARVEKRELMKENNLOGLVTSPLYNLRYLTNFTGTTGLAMLT 52
 86 GQPMWRSPGDNITTYTDMRDNFRAVROLTTGAKRIGIFDHY-----NLDF 133
 53 LDKAF---FVDPFRTE-----QAABQATGFTYIKNTGHPDEVDADLAERLQNDLAF 102
 134 RRO-----LEBALPGVEFVDISQPSMMWRITKSLEOKLIREGARVCDVGAACA 184
 103 ERTQVSPADYSLLEILP---CELVPMVGLIELEFVDEEVEVSIIEKACAIADOGFAFVLE 161
 165 AIKAGVPEHEVAIATNTMIREIAKSPFVELMDTWTFQSGINTDGAHNPVTNRYVSG 244

162 MIKPGMEIEVANQLDFPMNSKASGVSEFTIV-----ASGLRSAMPBGVAHSKVIKRG 215
 245 DILSLNTPMIFGYTTALERTLFCDHVDASLIDIEKNVAVHRGELIKPGARCKDIAI 304
 216 ELITLDFGCGYEGVSDMTRTFAIGSIOPKLEIYDYLEAQKLVABAKGEL---TGI 271
 305 ELNMYREMDLLKTRSPGYGHSFGVLYCHYGRAGVELREDIDTELKPGMYVSM 358
 272 QLDIARD---HIASIGDAPG---HSTHGIGLEIHEGPNNVSFRADKQFVPGONTITD 324
 359 EPMVLEPGMGAGYREHDLIYGEDEGAENITFP 394
 325 EPGIYL---FGIGVRIEDDLITAGGNVLTAP 356

RESULT 7
 US-09-107-532A-4292
 ; Sequence 4292, Application US/09107532A
 ; Patent No. 6583275

GENERAL INFORMATION:
 APPLICANT: Lynn A Doucette-Stamm and David Bush
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
 NUMBER OF SEQUENCES: 7310
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: GENOME THERAPEUTICS CORPORATION
 STREET: 100 Beaver Street
 CITY: Waltham
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02154

COMPUTER READABLE FORM:
 MEDIUM TYPE: CD-ROM ISO9660

OPERATING SYSTEM: <Unknown>
 SOFTWARE: ASCII

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/107,532A
 FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/085,598
 FILING DATE: 14 May 1998

APPLICATION NUMBER: 60/051571
 FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:
 NAME: Atinello, Pamela Deneke

REGISTRATION NUMBER: 40,489
 REFERENCE/DOCKET NUMBER: GTC-0112

TELECOMMUNICATION INFORMATION:
 TELEPHONE: (781) 893-5007
 TELEFAX: (781) 893-8277

INFORMATION FOR SEQ ID NO: 4292:
 LENGTH: 366 amino acids

SEQUENCE CHARACTERISTICS:
 TYPE: amino acid

TOPOLOGY: linear
 MOLECULE TYPE: protein

HYPOTHETICAL: YES
 ORIGINAL SOURCE:
 ORGANISM: Enterococcus faecium

FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (8) LOCATION 1...366

SEQUENCE DESCRIPTION: SEQ ID NO: 4292:
 US-09-107-532A-4292

Query Match 10.0%; Score 219; DB 4; Length 366;
 Best Local Similarity 23.5%; Pred. No. 6.7e-15;
 Matches 96; Conservative 63; Mismatches 177; Indels 72; Gaps 12;

11 WHNGEKDYSPSDAEMTRNDVRGMMAKNVDAALFTSYHCINYSGLYCYGRKGM 70
 1 WKACKERGSFL---MRLRVEKLRKKQENLDSFLVTSPLYNLRYLTNFTGTTGLA 53

CURRENT FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/107,433
PRIOR FILING DATE: 1998-06-30
PRIOR APPLICATION NUMBER: US 60/085,131
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: US 60/051,553
PRIOR FILING DATE: 1997-07-02
NUMBER OF SEQ ID NOS: 5322
SEQ ID NO 3285
LENGTH: 360
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-583-110-3285

Query Match 8.6%; Score 188.5; DB 4; Length 360;
Best Local Similarity 23.9%; Pred. No. 1.3e-11;
Matches 76; Conservative 50; Mismatches 139; Indels 53; Gaps 11;

QY 84 IDGGPWRSPGNDITTYDWRDNFRAVROLTTGAKRIGIEFDHNLDFRQLBEALPG 143
DB 76 VDSNPWK-----IKHALPOL--DFKRAVAFEDNLITKHGLKTVET 118
QY 144 VEFVDISOPSMWRTIKSLBE-OKLIREGARVCDVGAACAATKAGV-----PEHEV 195
DB 119 AEFNDLTRIQRMRLIKSADEVQKM-----VAGLYADKAVHVGFDNISLDTETDI 170
QY 196 AIAATNAMIRE-IKSPFVELMDTWTWFGSINTDGAH-PYTKRIYQSGDILSLNTP 253
DB 171 IAQIDPAMKREGVEMSPD-----TWVLTDGNANPHGIPPAANK-VENDLALLFDLGV 221
QY 254 MIFGYTALERTLPCDHVDASLDIWEKNVAVHRGELIKPGARCKOIAELNEMVREM 313
DB 222 LVNGYASDMTRIVAVGKPDQFKDITNLTLBAQNALDFIKGVTAAHEVDRAARVIEKA 281
QY 314 DLKTRSPFGYSHSFVLCHYRGREAGVELREDIDTELKPGVNVSMPEVMLEPGMGAG 373
DB 282 GGYEYFNRLHGHGIMDVHEF-----PSIMEGNDVIEGMCFSVEPGIYI----PGKVG 332
QY 374 YREHDILVGEDGAENIT 391
DB 333 VRIEDCGVVTXKGFDLFT 350

RESULT 11
US-09-107-433-3390
Sequence 3390, Application US/09107433

PATENT NO. 6800744
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGNO
THERAPEUTICS

NUMBER OF SEQUENCES: 5206
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION

STREET: 100 Beaver Street
CITY: Waltham

STATE: Massachusetts
COUNTRY: USA

ZIP: 02354
COMPUTER READABLE FORM:

MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: <Unknown>

OPERATING SYSTEM: <Unknown>
SOFTWARE: <Unknown>

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,433

FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/ 085131
FILING DATE: May 12, 1998

APPLICATION NUMBER: 60/051553
FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Ariniello, Pamela Deneke

REGISTRATION NUMBER: 40,489

REFERENCE/DOCKET NUMBER: GTC-011

TELECOMMUNICATION INFORMATION:

TELEPHONE: (781) 893-5007

TELEFAX: (781) 893-8277

INFORMATION FOR SEQ ID NO: 3390:

SEQUENCE CHARACTERISTICS:

LENGTH: 361 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: YES

ORIGINAL SOURCE:

ORGANISM: Streptococcus pneumoniae

FEATURE:

NAME/KEY: misc feature

LOCATION: (B) LOCATION 1..361

SEQUENCE DESCRIPTION: SEQ ID NO: 3390:

US-09-107-433-3390

Query Match 8.6%; Score 188.5; DB 4; Length 361;
Best Local Similarity 23.9%; Pred. No. 1.3e-11;
Matches 76; Conservative 50; Mismatches 139; Indels 53; Gaps 11;

QY 84 IDGGPWRSPGNDITTYDWRDNFRAVROLTTGAKRIGIEFDHNLDFRQLBEALPG 143
DB 77 VDSNPWK-----IKHALPOL--DFKRAVAFEDNLITKHGLKTVET 119
QY 144 VEFVDISOPSMWRTIKSLBE-OKLIREGARVCDVGAACAATKAGV-----PEHEV 195
DB 120 AEFNDLTRIQRMRLIKSADEVQKM-----VAGLYADKAVHVGFDNISLDTETDI 171
QY 196 AIAATNAMIRE-IKSPFVELMDTWTWFGSINTDGAH-PYTKRIYQSGDILSLNTP 253
DB 172 IAQIDPAMKREGVEMSPD-----TWVLTDGNANPHGIPPAANK-VENDLALLFDLGV 222
QY 254 MIFGYTALERTLPCDHVDASLDIWEKNVAVHRGELIKPGARCKOIAELNEMVREM 313
DB 223 LVNGYASDMTRIVAVGKPDQFKDITNLTLBAQNALDFIKGVTAAHEVDRAARVIEKA 282
QY 314 DLKTRSPFGYSHSFVLCHYRGREAGVELREDIDTELKPGVNVSMPEVMLEPGMGAG 373
DB 283 GGYEYFNRLHGHGIMDVHEF-----PSIMEGNDVIEGMCFSVEPGIYI----PGKVG 333
QY 374 YREHDILVGEDGAENIT 391
DB 334 VRIEDCGVVTXKGFDLFT 351

RESULT 12

US-09-602-777A-94
Sequence 94, Application US/09602777A

PATENT NO. 6831165
GENERAL INFORMATION:

APPLICANT: Pompeius, Markus

APPLICANT: Kroger, Burkhard

APPLICANT: Schroder, Hartwig

APPLICANT: Zelder, Oskar

APPLICANT: Haberer, Gregor

TITLE OF INVENTION: CORNBACTERIUM GLUTAMINUM GENES ENCODING PROTEINS

TITLE OF INVENTION: INVOLVED IN HOMEOSTASIS AND ADAPTATION

FILE REFERENCE: BGI-128CP
CURRENT APPLICATION NUMBER: US/09/602,777A
PRIOR FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: US 60/141031
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: DE 19931636.8
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19932125.6
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932126.4

PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932127.2
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932128.0
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932129.9
PRIOR FILING DATE: 1999-07-19
PRIOR APPLICATION NUMBER: DE 19932226.0
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932920.6
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19932922.2
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19932924.9
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19932930.3
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19932933.8
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19932935.4
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19932973.7
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19933002.6
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19933003.4
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19933005.0
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19933006.9
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19941378.9
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: DE 19941379.7
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: DE 19941390.8
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: DE 19941391.6
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: DE 19942088.2
PRIOR FILING DATE: 1999-09-03
NUMBER OF SEQ ID NOS: 442
SEQ ID NO 94
LENGTH: 363
TYPE: PRP
ORGANISM: Corynebacterium glutamicum
US-09-602-777A-94

Query Match 8.6%; Score 187.5; DB 4; Length 363;
Best Local Similarity 22.2%; Pred. No. 1.7e-11;
Matches 91; Conservative 56; Mismatches 164; Indels 99; Gaps 12;
QY 22 SPAEMTRNDVRCGMAKKNVDAALFTSHCINYYSGWL-----YCYFGR 66
DB 4 AOTRPAIRRALAAKAAQRIIDILVTSPIHAYILSGFGSCALIVNKDLSAQITDGR 63
QY 67 KYGMV-----IDHNNATTISAGIDGQFWRRSFGDNITYTDRDNFYRAVROLT 116
DB 64 YTTQIAEVPDIEALIERASATTLAQVE----- 92
QY 117 TCAKRIIGIFPDHVNLDPRROLEAL-PGYEVDISOPSMWRTIKSLEOKLIREGARVC 175
DB 93 -OPRRAIRBAQTTLDQLDSLEATOEDVELIPVS-----GVVESIRLTKOSFELDRLR 145
QY 176 DVGGA-----CAAIRKAGVBEHVAIATNAMIIRIAKSPFVELMDTWTFQSGI 227
DB 146 DVAALASQAFEDLLAAGEIABEGSEROVA-ADLEYRRLIGAERPFDD-----TIYASGP 199
QY 228 NTDGANPVTNRIVOSGDILSLNTFPMIGYATALERTLFCDDVDDASLDIWEKNAVAR 287
DB 200 NAKKPHHGAGDRIILQNGDLVTIDFGAARGFNSDMRTITLWGEAGFEBAIYDIYLRSQL 259

QY 288 RLELEIKGACRCKDIAIEINEMREMDLKYRSFGYHSFGVLCHYYGAEVLEDEDID 347
DB 260 AGEVAAYGANLFDIDAACKRIED-----AGYGEYF---VHSTGIGIGLEVHEAPS 308
QY 348 TE-----LKPGMVVSMEDPVMVLPFGMGAGCYREHDIIVGEDGAENIT 391
DB 309 ASKTSQVLETSSTLIEPGIY---VPGKGVRIEDTLITNSGAPEIIT 354
RESULT 13
US-09-583-110-4793
Sequence 4793, Application US/09583110
Patent No. 6699703
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al.
TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
FILE REFERENCE: PATH00-07A
CURRENT APPLICATION NUMBER: US/09/583,110
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/107,433
PRIOR FILING DATE: 1998-06-30
PRIOR APPLICATION NUMBER: US 60/085,131
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: US 60/051,553
PRIOR FILING DATE: 1997-07-02
NUMBER OF SEQ ID NOS: 5322
SEQ ID NO 4793
LENGTH: 353
TYPE: PRP
ORGANISM: Streptococcus pneumoniae
US-09-583-110-4793

Query Match 8.4%; Score 183; DB 4; Length 353;
Best Local Similarity 24.4%; Pred. No. 5.1e-11;
Matches 71; Conservative 55; Mismatches 121; Indels 44; Gaps 11;
QY 118 GAKRIGIFPDHVNLDPRROLEALPGVEFVDISOPSMWRTIKSLEOKLIREGARVCV 177
DB 87 GLTRIGFE-DEISVYTHRMQAAFGGLDFPOTQVPEGIRMTKDEEIAIRKACISIQ 145
QY 178 GGAACAAMIKAGVBEHVAIATNAMIIRIAKSPFVELMDTWTFQSGINTGAINPVT 237
DB 146 AFRDALDPFKPKTEIEIA-NFLDPRMELASG-----LSPTDILASGINSKPHAHM 199
QY 238 NRIVOSGDILSLNTFPMIFGYTTA-LERTLFCDDVDDASLDIWEKNAVARHKGLEIKFG 296
DB 200 HKPVEIGEAITWD-FGCIYDHYVSDMTRITLYGHVSDQAEIYNTVLYKANQALIDQAKG 258
QY 297 AAC-----KDIAIENEMREMDLKKRSFGYHSFGVLCHYYGAEVLEDEDID 345
DB 259 LGFRDFDKIPRDIIEA-----GYGDYF---THGIGHGIGLDIHEEPPYS 300
QY 346 -LDTE-LKRGVVSMEPMVLPFGMGAGCYREHDIIVGEDGAENITGFP 394
DB 301 QSTETIKGMALTDPEGIY---EGKGVAIEDDILITETGCELLTAP 347
RESULT 14
US-09-107-433-3157
Sequence 3157, Application US/09107433
Patent No. 6800744
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
THERAPEUTICS
NUMBER OF SEQUENCES: 5206
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham

```

STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: <Unknown>
OPERATING SYSTEM: <Unknown>
SOFTWARE: <Unknown>
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,433
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085131
FILING DATE: May 12, 1998
APPLICATION NUMBER: 60/051553
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Arinello, Pamela Dencke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-011
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 3157:
SEQUENCE CHARACTERISTICS:
LENGTH: 358 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Streptococcus pneumoniae
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...358
SEQUENCE DESCRIPTION: SEQ ID NO: 3157:
US-09-107-433-3157

Query Match      8.4%; Score 183; DB 4; Length 358;
Best Local Similarity 24.4%; Pred. No. 5.2e-11;
Matches 71; Conservative 55; Mismatches 121; Indels 44; Gaps 11;

QY 118 GAKRIGIEPDHVNLDPRQLEALPGVEFVDISQPSMMWRTIKSLSEQKLIREGARVCDV 177
   ||| ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 92 GLTRIGFE-DEISVSYYHMOAFAGLDFPQTQFVGLRMIXDEALIAIKRACSIDQ 150
   ||| ||| : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 178 GGAACAATKAGVEHEVAIATTNAMIRESLSPFVELMDTWTWFOGINTDGAHPVT 237
   ||| ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 151 AFRDALDFIKPKTELEIA-NFLDFRRELGAAG-----LSFDTLASGINSKPHAPM 204
   ||| ||| : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 238 NRIVSGDILLNTPPMIFGYTA-LERTLPCDHVDDASLDIWEKNVAVHRGELIKPG 296
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 205 HKPVELGEALIMD-FCGLDYHVSMDTTRITLYGHVSDQEAELYNVILKANQALIDQAKAG 263
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 297 ARC-----KQIAIFLNEMREMDLTKRSFGYGHSPGLCHYYREAGVELRED---- 345
   ||| ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 264 LGFRDFDKLPRDIIIEA-----GYGDYF---TIGIGHGIGLDIHEEPYFS 305
   ||| ||| : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 346 -IDTE-LKPGVVSMPVMLPEGMFGAGYREHDLIVGEDGAENITGFP 394
   ||| ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 306 QTSSTIKTGMLIDEPGIYI---EGKGVRIEDDILITETGCCHLTAP 352
   ||| ||| : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 15
US-09-861-451A-36
Sequence 36, Application US/09861451A
Patent No. 675316
GENERAL INFORMATION:
APPLICANT: Commonwealth Scientific & Industrial Research Orga
TITLE OF INVENTION: Methods of Identifying Antigen Gene Sequences
FILE REFERENCE: FP34033/01
CURRENT APPLICATION NUMBER: US/09/861,451A
CURRENT FILING DATE: 2001-05-21

```

```

PRIOR APPLICATION NUMBER: PP7273
PRIOR FILING DATE: 1998-11-20
NUMBER OF SEQ ID NOS: 84
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 36
LENGTH: 299
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Deduced protein
US-09-861-451A-36

Query Match      8.3%; Score 181.5; DB 4; Length 299;
Best Local Similarity 22.5%; Pred. No. 5.7e-11;
Matches 61; Conservative 49; Mismatches 132; Indels 29; Gaps 6;

QY 127 DHVNDLPRQLEALPGVEFVDISQPSMMWRTIKSLSEQKLIREGARVCDVGAACAAT 186
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 46 EYLITDQDKIRSWFPNADPVYKQ--KQLPRIIKTEELIKNIKAVEISLAIVNKIPKI 103
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 187 KAGVEHEVAIATTNAMIRESLSPFVELMDTWTWFOGINTDGAHPVTNRIVQSGDI 246
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 104 KPGMTEKSIDVNLNQMILGAEKESFDSIIAT-----GNSAMPHRASSETIILDNDL 157
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 247 LSLNTPPMIFGYTALERTLPCDHVDDASLDIWEKNVAVHRGELIKPGARCKQIATL 306
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 158 LKIDGALFNQCADITFTSYIGQISEKDLIELVKCAELIGKKVAPGVKASIDILAC 217
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 307 NEMVREMDLTKRSPGYGHSPGLCHYYREAGVELRE-----DIDTELKPGVVSMEP 360
   ||| ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 218 RNFTIEQ-----GYKVF---IHSTGHGVGIDIHLLPVVSSQTILPEGVITVER 266
   ||| ||| : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 361 WVMLEPGMFGAGYREHDLIVGEDGAENIT 391
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 267 GIYI---PGLGARIEDVVLVTESGFRTLS 293
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Search completed: July 8, 2005, 01:36:46
 Job time : 29 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 8, 2005, 01:35:47 ; Search time 77 Seconds

(without alignment)
2026.980 Million cell updates/sec

Title: US-10-807-228a-1

Perfect score: 2190

Sequence: 1 MTDDMLHMKHNGEKDYSP.....DGAENITGPPGPHNITRN 404

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1726216 seqs, 386330316 residues

Total number of hits satisfying chosen parameters: 1726216

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2078	94.9	404	US-10-251-078-2	Sequence 2, Appl
2	2078	94.9	404	US-10-251-078-18	Sequence 18, Appl
3	2068	94.4	404	US-10-251-078-10	Sequence 10, Appl
4	2068	94.4	404	US-10-251-078-22	Sequence 22, Appl
5	2066	94.3	404	US-10-251-078-16	Sequence 16, Appl
6	2064	94.2	404	US-10-251-078-12	Sequence 12, Appl
7	2062	94.2	404	US-10-251-078-14	Sequence 14, Appl
8	2058	94.0	404	US-10-251-078-20	Sequence 20, Appl
9	248.5	10.5	354	US-09-815-242-10645	Sequence 10645, A
10	229.5	10.5	354	US-09-815-242-10645	Sequence 10645, A
11	229	10.5	361	US-09-815-242-10235	Sequence 10235, A

12	203.5	9.3	381	14	US-10-156-761-9699	Sequence 9699, Ap
13	200	9.1	350	9	US-09-815-242-5287	Sequence 5287, Ap
14	200	9.1	353	9	US-09-815-242-12580	Sequence 12580, A
15	194.5	8.9	353	15	US-10-282-122A-70734	Sequence 70734, A
16	190.5	8.7	360	17	US-10-472-928-3260	Sequence 3260, Ap
17	189	8.6	353	9	US-09-815-242-13193	Sequence 13193, A
18	188.5	8.6	361	18	US-10-617-320-3390	Sequence 3390, Ap
19	187.5	8.6	363	9	US-09-738-626-5290	Sequence 5290, Ap
20	187.5	8.6	368	14	US-10-156-761-14386	Sequence 14386, A
21	183	8.4	353	17	US-10-472-928-184	Sequence 184, Ap
22	183	8.4	358	18	US-10-617-320-3157	Sequence 3157, Ap
23	181.5	8.3	293	9	US-09-861-451A-36	Sequence 36, Appl
24	179	8.2	243	10	US-09-877-843-94	Sequence 94, Appl
25	171	7.8	355	15	US-10-289-762-871	Sequence 871, Appl
26	166	7.6	240	15	US-10-042-865-189	Sequence 189, Appl
27	162.5	7.4	375	16	US-10-482-706-141	Sequence 141, Appl
28	160.5	7.3	253	15	US-10-289-867-10	Sequence 10, Appl
29	160.5	7.3	253	15	US-10-350-516-4	Sequence 4, Appl
30	160.5	7.3	397	15	US-10-424-599-247413	Sequence 247413, A
31	160	7.3	352	15	US-10-282-122A-71551	Sequence 71551, A
32	158.5	7.2	350	16	US-10-425-115-369034	Sequence 369034, A
33	156.5	7.1	306	15	US-10-425-114-58168	Sequence 58168, A
34	155	7.1	428	15	US-10-369-493-245	Sequence 245, Appl
35	152.5	7.0	378	15	US-10-369-493-16426	Sequence 16426, A
36	151.5	6.9	417	15	US-10-369-493-10239	Sequence 10239, A
37	151	6.9	391	16	US-10-425-115-299292	Sequence 299292, A
38	146.5	6.7	437	15	US-10-369-493-21222	Sequence 21222, A
39	145	6.6	356	15	US-10-424-599-173546	Sequence 173546, A
40	143.5	6.6	369	15	US-10-042-865-186	Sequence 186, Appl
41	143	6.5	379	9	US-09-738-626-5142	Sequence 5142, Ap
42	142.5	6.5	369	15	US-10-042-865-187	Sequence 187, Appl
43	142	6.5	274	15	US-10-299-867-11	Sequence 11, Appl
44	141	6.4	432	16	US-10-739-930-9633	Sequence 9633, Ap
45	140.5	6.4	248	15	US-10-299-867-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1
US-10-251-078-2
; Sequence 2, Application US/10251078
; Publication No. US200301119084A1
; GENERAL INFORMATION:
; APPLICANT: Roche Diagnostics GmbH
; APPLICANT: Shao, Zhixin
; APPLICANT: Schmuck, Rainer
; APPLICANT: Kratzsch, Peter
; APPLICANT: Kemlises, Janet
; APPLICANT: Weisner, Harald
; TITLE OF INVENTION: Variants of an Erwinia-type creatinase
; FILE REFERENCE: 20981 EP
; CURRENT APPLICATION NUMBER: US/10/251.078
; CURRENT FILING DATE: 2002-09-20
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 404
; TYPE: PRT
; ORGANISM: Erwinia sp.
US-10-251-078-2

Query Match 94.9%, Score 2078, DB 14, Length 404;
Best Local Similarity 94.3%, Pred. No. 4.1e-195;
Matches 381, Conservative 14, Mismatches 9, Indels 0, Gaps 0;

QY 1 MTDDMLHMKHNGEKDYSPSDAEMTRQNDVGRMAKNNVDAALFTSYHCINYSGL 60
DB 1 MTDDMLHMKHNGEKDYSPSDAEMTRQNDVGRMAKNNVDAALFTSYHCINYSGL 60
QY 61 YCYGRKXGWIIDHNATTTISAGIDGQPMRRSPFDNITTYTDWRDNNFYRAVRLTGA 120
DB 61 YCYGRKXGWIIDHNATTTISAGIDGQPMRRSPFDNITTYTDWRDNNFYRAVRLTGA 120

QY 121 RIGIEFDHVNIDFRROLEBALPGVEFVDISOPSMWRTIKSLIEOKLIREGARVCDVGA 180
DB 121 RIGIEFDHVNIDFRRTLEBALPGVEFVDIGOPSMWRTIKSLIEOKLIREGARICDVGA 180
QY 181 ACAAIKAGVPEHEVAIAITTNAMIREIAKSPFEVLDWTWFOGINTDGAHNPVTNR 240
DB 181 ACVAAVKAGVPEHEVAIAITTNAMIREIAKSPFEVLDWTWFOGINTDGAHNPVTNR 240
QY 241 VOSGDIISLNTFPMIFGYTTALERTLFCDHVDASLDIMEKNVAVHRGLELIKRGARCK 300
DB 241 VOSGDIISLNTFPMIFGYTTALERTLFCDHVDASLDIMEKNVAVHRGLELIKRGARCK 300
QY 301 DIAELNEMTYREMDLTKRSPGYGHSFGVLSHYGREGAVELREDIDVTLOPGMVVSMEP 360
DB 301 DIAELNEMTYREMDLTKRSPGYGHSFGVLSHYGREGAVELREDIDVTLOPGMVVSMEP 360
QY 361 MWMLPEGAPGAGYREHDLIVGEDGABNTGPFPGPEHNITRN 404
DB 361 MWMLPEGAPGAGYREHDLIVGEDGABNTGPFPGPEHNITRN 404

RESULT 2
US-10-251-078-18

; Sequence 18, Application US/10251078
; Publication No. US20030119084A1
; GENERAL INFORMATION:
; APPLICANT: Roche Diagnostics GmbH
; APPLICANT: Shao, Zhixin
; APPLICANT: Schmock, Rainer
; APPLICANT: Kratzsch, Peter
; APPLICANT: Kenkies, Janet
; APPLICANT: Weisner, Harald
; TITLE OF INVENTION: Variants of an Erwinia-type creatinase
; FILE REFERENCE: 20981 BP
; CURRENT APPLICATION NUMBER: US/10/251,078
; CURRENT FILING DATE: 2002-09-20
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 404
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: variant C19C2
US-10-251-078-18

Query Match 94.9%; Score 2078; DB 14; Length 404;
Best Local Similarity 94.1%; Pred. No. 4, 1e-195;
Matches 380; Conservative 14; Mismatches 10; Indels 0; Gaps 0;

QY 1 MTDDMLHVMKHNKNGEKDYSFSDAEMTRRONDVRGMANKNVDAALFTSYHCINYYSGML 60
DB 1 MTDDMLHVMKHNKNGEKDYSFSDAEMTRRONDVRGMANKNVDAALFTSYHCINYYSGML 60
QY 61 YCYFGRKXGMYIDDDHATTISAGIDGQPMWRSPGDNITTYDMRDNFYQAVROLTPGAR 120
DB 61 YCYFGRKXGMYIDDDHATTISAGIDGQPMWRSPGDNITTYDMRDNFYQAVROLTPGAR 120
QY 121 RIGIEFDHVNIDFRROLEBALPGVEFVDISOPSMWRTIKSLIEOKLIREGARVCDVGA 180
DB 121 RIGIEFDHVNIDFRRTLEBALPGVEFVDIGOPSMWRTIKSLIEOKLIREGARICDVGA 180
QY 181 ACAAIKAGVPEHEVAIAITTNAMIREIAKSPFEVLDWTWFOGINTDGAHNPVTNR 240
DB 181 ACVAAVKAGVPEHEVAIAITTNAMIREIAKSPFEVLDWTWFOGINTDGAHNPVTNR 240
QY 241 VOSGDIISLNTFPMIFGYTTALERTLFCDHVDASLDIMEKNVAVHRGLELIKRGARCK 300
DB 241 VOSGDIISLNTFPMIFGYTTALERTLFCDHVDASLDIMEKNVAVHRGLELIKRGARCK 300
QY 301 DIAELNEMTYREMDLTKRSPGYGHSFGVLSHYGREGAVELREDIDVTLOPGMVVSMEP 360
DB 301 DIAELNEMTYREMDLTKRSPGYGHSFGVLSHYGREGAVELREDIDVTLOPGMVVSMEP 360

DB 301 DIAELNEMTYREMDLTKRSPGYGHSFGVLSHYGREGAVELREDIDVTLOPGMVVSMEP 360
QY 361 MWMLPEGAPGAGYREHDLIVGEDGABNTGPFPGPEHNITRN 404
DB 361 MWMLPEGAPGAGYREHDLIVGEDGABNTGPFPGPEHNITRN 404

RESULT 3
US-10-251-078-10

; Sequence 10, Application US/10251078
; Publication No. US20030119084A1
; GENERAL INFORMATION:
; APPLICANT: Roche Diagnostics GmbH
; APPLICANT: Shao, Zhixin
; APPLICANT: Schmock, Rainer
; APPLICANT: Kratzsch, Peter
; APPLICANT: Kenkies, Janet
; APPLICANT: Weisner, Harald
; TITLE OF INVENTION: Variants of an Erwinia-type creatinase
; FILE REFERENCE: 20981 BP
; CURRENT APPLICATION NUMBER: US/10/251,078
; CURRENT FILING DATE: 2002-09-20
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 404
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: variant C11m24
US-10-251-078-10

Query Match 94.4%; Score 2068; DB 14; Length 404;
Best Local Similarity 93.8%; Pred. No. 3, 9e-194;
Matches 379; Conservative 15; Mismatches 10; Indels 0; Gaps 0;

QY 1 MTDDMLHVMKHNKNGEKDYSFSDAEMTRRONDVRGMANKNVDAALFTSYHCINYYSGML 60
DB 1 MTDDMLHVMKHNKNGEKDYSFSDAEMTRRONDVRGMANKNVDAALFTSYHCINYYSGML 60
QY 61 YCYFGRKXGMYIDDDHATTISAGIDGQPMWRSPGDNITTYDMRDNFYQAVROLTPGAR 120
DB 61 YCYFGRKXGMYIDDDHATTISAGIDGQPMWRSPGDNITTYDMRDNFYQAVROLTPGAR 120
QY 121 RIGIEFDHVNIDFRROLEBALPGVEFVDISOPSMWRTIKSLIEOKLIREGARVCDVGA 180
DB 121 RIGIEFDHVNIDFRRTLEBALPGVEFVDIGOPSMWRTIKSLIEOKLIREGARICDVGA 180
QY 181 ACAAIKAGVPEHEVAIAITTNAMIREIAKSPFEVLDWTWFOGINTDGAHNPVTNR 240
DB 181 ACVAAVKAGVPEHEVAIAITTNAMIREIAKSPFEVLDWTWFOGINTDGAHNPVTNR 240
QY 241 VOSGDIISLNTFPMIFGYTTALERTLFCDHVDASLDIMEKNVAVHRGLELIKRGARCK 300
DB 241 VOSGDIISLNTFPMIFGYTTALERTLFCDHVDASLDIMEKNVAVHRGLELIKRGARCK 300
QY 301 DIAELNEMTYREMDLTKRSPGYGHSFGVLSHYGREGAVELREDIDVTLOPGMVVSMEP 360
DB 301 DIAELNEMTYREMDLTKRSPGYGHSFGVLSHYGREGAVELREDIDVTLOPGMVVSMEP 360
QY 361 MWMLPEGAPGAGYREHDLIVGEDGABNTGPFPGPEHNITRN 404
DB 361 MWMLPEGAPGAGYREHDLIVGEDGABNTGPFPGPEHNITRN 404

RESULT 4
US-10-251-078-22

; Sequence 22, Application US/10251078
; Publication No. US20030119084A1
; GENERAL INFORMATION:
; APPLICANT: Roche Diagnostics GmbH
; APPLICANT: Shao, Zhixin
; APPLICANT: Schmock, Rainer

APPLICANT: Kratzsch, Peter
APPLICANT: Kenkies, Janet
TITLE OF INVENTION: Variants of an Erwinia-type creatinase
FILE REFERENCE: 20981 EP
CURRENT APPLICATION NUMBER: US/10/251,078
CURRENT FILING DATE: 2002-09-20
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PatentIn version 3.1
SEQ ID NO 22
LENGTH: 404
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:variant CT2m7
US-10-251-078-22

Query Match 94.4%; Score 2068; DB 14; Length 404;
Best Local Similarity 93.6%; Pred. No. 3.9e-194;
Matches 378; Conservative 15; Mismatches 11; Indels 0; Gaps 0;

QY 1 MTDDMLHMKHNGEKDYSPSDAEMTRQNDVRGMAKNVDAALFTSYHCINYSGL 60
DB 1 MTDDMLHMKHNGEKDYSPSDAEMTRQNDVRGMAKNVDAALFTSYHCINYSGL 60
QY YCYFGRKYGWIDHNNATTISAGIDGQPMWRSPGDNITTYDMRDNFYRAVROLTPGAK 120
DB 61 YCYFGRKYGWIDHNNATTISAGIDGQPMWRSPGDNITTYDMRDNFYRAVROLTPGAK 120
QY 121 RIGIEFDHVNLDPRQLEALPGVEFVDISQPSMMRTIKSLSEOKLIREGARVCDVGA 180
DB 121 RIGIEFDHVNLDPRQLEALPGVEFVDISQPSMMRTIKSLSEOKLIREGARVCDVGA 180
QY 181 ACAAIKAGVPEHEVAIATTNATIREIAKSPFVELMDTWTWFOGINTDGAHNPVTNRI 240
DB 181 ACVAAVKAGVPEHEVAIATTNATIREIAKSPFVELMDTWTWFOGINTDGAHNPVTNRI 240
QY 241 VOSGDILSLNTPPMIFGYTTALERTLPCDHVDASLDITWKNVAHRRGLIELIKGARCK 300
DB 241 VOSGDILSLNTPPMIFGYTTALERTLPCDHVDASLDITWKNVAHRRGLIELIKGARCK 300
QY 301 DIAELNEMREMDLTKRSPGYGHSFGVLCHYGRAGAVELREDIDTELKPGVVSMEP 360
DB 301 DIAELNEMREMDLTKRSPGYGHSFGVLCHYGRAGAVELREDIDTELKPGVVSMEP 360
QY 361 MVMLEPGMFGAGGYREHDLIVGEDGAENITGPPGPEHNIIRN 404
DB 361 MVMLEPGMFGAGGYREHDLIVGEDGAENITGPPGPEHNIIRN 404

RESULT 5
US-10-251-078-16
Sequence 16, Application US/10251078
Publication No. US20030119084A1
GENERAL INFORMATION:
APPLICANT: Roche Diagnostics GmbH
APPLICANT: Shao, Zhixin
APPLICANT: Schmuck, Rainer
APPLICANT: Kratzsch, Peter
APPLICANT: Kenkies, Janet
APPLICANT: Weisner, Harald
TITLE OF INVENTION: Variants of an Erwinia-type creatinase
FILE REFERENCE: 20981 EP
CURRENT APPLICATION NUMBER: US/10/251,078
CURRENT FILING DATE: 2002-09-20
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PatentIn version 3.1
SEQ ID NO 16
LENGTH: 404
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:variant CT2m8

US-10-251-078-16
Query Match 94.3%; Score 2066; DB 14; Length 404;
Best Local Similarity 93.6%; Pred. No. 6.2e-194;
Matches 378; Conservative 16; Mismatches 10; Indels 0; Gaps 0;

QY 1 MTDDMLHMKHNGEKDYSPSDAEMTRQNDVRGMAKNVDAALFTSYHCINYSGL 60
DB 1 MTDDMLHMKHNGEKDYSPSDAEMTRQNDVRGMAKNVDAALFTSYHCINYSGL 60
QY 61 YCYFGRKYGWIDHNNATTISAGIDGQPMWRSPGDNITTYDMRDNFYRAVROLTPGAK 120
DB 61 YCYFGRKYGWIDHNNATTISAGIDGQPMWRSPGDNITTYDMRDNFYRAVROLTPGAK 120
QY 121 RIGIEFDHVNLDPRQLEALPGVEFVDISQPSMMRTIKSLSEOKLIREGARVCDVGA 180
DB 121 RIGIEFDHVNLDPRQLEALPGVEFVDISQPSMMRTIKSLSEOKLIREGARVCDVGA 180
QY 181 ACAAIKAGVPEHEVAIATTNATIREIAKSPFVELMDTWTWFOGINTDGAHNPVTNRI 240
DB 181 ACVAAVKAGVPEHEVAIATTNATIREIAKSPFVELMDTWTWFOGINTDGAHNPVTNRI 240
QY 241 VOSGDILSLNTPPMIFGYTTALERTLPCDHVDASLDITWKNVAHRRGLIELIKGARCK 300
DB 241 VOSGDILSLNTPPMIFGYTTALERTLPCDHVDASLDITWKNVAHRRGLIELIKGARCK 300
QY 301 DIAELNEMREMDLTKRSPGYGHSFGVLCHYGRAGAVELREDIDTELKPGVVSMEP 360
DB 301 DIAELNEMREMDLTKRSPGYGHSFGVLCHYGRAGAVELREDIDTELKPGVVSMEP 360
QY 361 MVMLEPGMFGAGGYREHDLIVGEDGAENITGPPGPEHNIIRN 404
DB 361 MVMLEPGMFGAGGYREHDLIVGEDGAENITGPPGPEHNIIRN 404

RESULT 6
US-10-251-078-12
Sequence 12, Application US/10251078
Publication No. US20030119084A1
GENERAL INFORMATION:
APPLICANT: Roche Diagnostics GmbH
APPLICANT: Shao, Zhixin
APPLICANT: Schmuck, Rainer
APPLICANT: Kratzsch, Peter
APPLICANT: Kenkies, Janet
APPLICANT: Weisner, Harald
TITLE OF INVENTION: Variants of an Erwinia-type creatinase
FILE REFERENCE: 20981 EP
CURRENT APPLICATION NUMBER: US/10/251,078
CURRENT FILING DATE: 2002-09-20
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PatentIn version 3.1
SEQ ID NO 12
LENGTH: 404
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:variant CT2m9
US-10-251-078-12

Query Match 94.2%; Score 2064; DB 14; Length 404;
Best Local Similarity 93.6%; Pred. No. 9.7e-194;
Matches 378; Conservative 16; Mismatches 10; Indels 0; Gaps 0;

QY 1 MTDDMLHMKHNGEKDYSPSDAEMTRQNDVRGMAKNVDAALFTSYHCINYSGL 60
DB 1 MTDDMLHMKHNGEKDYSPSDAEMTRQNDVRGMAKNVDAALFTSYHCINYSGL 60
QY 61 YCYFGRKYGWIDHNNATTISAGIDGQPMWRSPGDNITTYDMRDNFYRAVROLTPGAK 120
DB 61 YCYFGRKYGWIDHNNATTISAGIDGQPMWRSPGDNITTYDMRDNFYRAVROLTPGAK 120
QY 121 RIGIEFDHVNLDPRQLEALPGVEFVDISQPSMMRTIKSLSEOKLIREGARVCDVGA 180
DB 121 RIGIEFDHVNLDPRQLEALPGVEFVDISQPSMMRTIKSLSEOKLIREGARVCDVGA 180

Db 121 RIGIEFDHVDLDFRRTLEALPGVEFVDIGQSPMMKRVKSLBEOKLIREBARICDVGA 180
Qy 181 ACAAIKAGVBEHEVALATTNAMIREIAKSPPEVLMDTWTWFOGINTDGAHPVTNRI 240
Db 181 ACVAAYKAGVBEHEVALATTNAMIREIAKSPPEVLMDTWTWFOGINTDGAHPVTNRI 240
Qy 241 VOSGDIIISLNTFPMIFGYTTLERTLFCDHVDASLDIWEKNVAVHRRGLELIRPGARCK 300
Db 241 VOSGDIIISLNTFPMIFGYTTLERTLFCDHVDASLDIWEKNVAVHRRGLELIRPGARCK 300
Qy 301 DIAELNEMRYEMDLKTRSPGYGSHFCVLCYYYGREGAVELREDIDITVLQPGMVVSWEP 360
Db 301 DIAELNEMRYEMDLKTRSPGYGSHFCVLCYYYGREGAVELREDIDITVLQPGMVVSWEP 360
Qy 361 MVMLEPGAPGAGYREHDLIVGEDGAMNTGPPGPEHNIRN 404
Db 361 MVMLEPGAPGAGYREHDLIVGEDGAMNTGPPGPEHNIRN 404

RESULT 7
; Sequence 14, Application US/10251078
; Publication No. US20030119084A1
; GENERAL INFORMATION:
; APPLICANT: Roche Diagnostics GmbH
; APPLICANT: Shao, Zhixin
; APPLICANT: Schmuck, Rainer
; APPLICANT: Kratzsch, Peter
; APPLICANT: Kenkies, Janet
; APPLICANT: Weisner, Harald
; TITLE OF INVENTION: Variants of an Erwinia-type creatinase
; FILE REFERENCE: 20981 EP
; CURRENT APPLICATION NUMBER: US/10/251,078
; CURRENT FILING DATE: 2002-09-20
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 404
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Description of Artificial Sequence:variant CT2m10
US-10-251-078-14

Query Match 94.2%; Score 2062; DB 14; Length 404;
Best Local Similarity 93.6%; Pred. No. 1.5e-193;
Matches 378; Conservative 15; Mismatches 11; Indels 0; Gaps 0;
Qy 1 MTDDMLHVMKMHNGEKDYSFSDAEMTRRQNDVGMKAKNVDAALFTSYHCINYYSGWL 60
Db 1 MTDDMLHVMKMHNGEKDYSFSDAEMTRRQNDVGMKAKNVDAALFTSYHCINYYSGWL 60
Qy 61 YCYGGRKYGWVIDDHAATTISAGIDGGQPMRRSFEDNITYTDMRRDNFYARVRLTGTAK 120
Db 61 YCYGGRKYGWVIDDHAATTISAGIDGGQPMRRSFEDNITYTDMRRDNFYARVRLTGTAK 120
Qy 121 RIGIEFDHVDLDFRRTLEALPGVEFVDISQSPMMKRTIKSLBEOKLIREGARVCDVGA 180
Db 121 RIGIEFDHVDLDFRRTLEALPGVEFVDISQSPMMKRTIKSLBEOKLIREGARVCDVGA 180
Qy 181 ACAAIKAGVBEHEVALATTNAMIREIAKSPPEVLMDTWTWFOGINTDGAHPVTNRI 240
Db 181 ACVAAYKAGVBEHEVALATTNAMIREIAKSPPEVLMDTWTWFOGINTDGAHPVTNRI 240
Qy 241 VOSGDIIISLNTFPMIFGYTTLERTLFCDHVDASLDIWEKNVAVHRRGLELIRPGARCK 300
Db 241 VOSGDIIISLNTFPMIFGYTTLERTLFCDHVDASLDIWEKNVAVHRRGLELIRPGARCK 300
Qy 301 DIAELNEMRYEMDLKTRSPGYGSHFCVLCYYYGREGAVELREDIDITVLQPGMVVSWEP 360
Db 301 DIAELNEMRYEMDLKTRSPGYGSHFCVLCYYYGREGAVELREDIDITVLQPGMVVSWEP 360

Qy 361 MVMLEPGAPGAGYREHDLIVGEDGAMNTGPPGPEHNIRN 404
Db 361 MVMLEPGAPGAGYREHDLIVGEDGAMNTGPPGPEHNIRN 404

RESULT 8
; Sequence 20, Application US/10251078
; Publication No. US20030119084A1
; GENERAL INFORMATION:
; APPLICANT: Roche Diagnostics GmbH
; APPLICANT: Shao, Zhixin
; APPLICANT: Schmuck, Rainer
; APPLICANT: Kratzsch, Peter
; APPLICANT: Kenkies, Janet
; APPLICANT: Weisner, Harald
; TITLE OF INVENTION: Variants of an Erwinia-type creatinase
; FILE REFERENCE: 20981 EP
; CURRENT APPLICATION NUMBER: US/10/251,078
; CURRENT FILING DATE: 2002-09-20
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 404
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Description of Artificial Sequence:variant CTsd2
US-10-251-078-20

Query Match 94.0%; Score 2058; DB 14; Length 404;
Best Local Similarity 93.3%; Pred. No. 3.8e-193;
Matches 377; Conservative 16; Mismatches 11; Indels 0; Gaps 0;
Qy 1 MTDDMLHVMKMHNGEKDYSFSDAEMTRRQNDVGMKAKNVDAALFTSYHCINYYSGWL 60
Db 1 MTDDMLHVMKMHNGEKDYSFSDAEMTRRQNDVGMKAKNVDAALFTSYHCINYYSGWL 60
Qy 61 YCYGGRKYGWVIDDHAATTISAGIDGGQPMRRSFEDNITYTDMRRDNFYARVRLTGTAK 120
Db 61 YCYGGRKYGWVIDDHAATTISAGIDGGQPMRRSFEDNITYTDMRRDNFYARVRLTGTAK 120
Qy 121 RIGIEFDHVDLDFRRTLEALPGVEFVDISQSPMMKRTIKSLBEOKLIREGARVCDVGA 180
Db 121 RIGIEFDHVDLDFRRTLEALPGVEFVDISQSPMMKRTIKSLBEOKLIREGARVCDVGA 180
Qy 181 ACAAIKAGVBEHEVALATTNAMIREIAKSPPEVLMDTWTWFOGINTDGAHPVTNRI 240
Db 181 ACVAAYKAGVBEHEVALATTNAMIREIAKSPPEVLMDTWTWFOGINTDGAHPVTNRI 240
Qy 241 VOSGDIIISLNTFPMIFGYTTLERTLFCDHVDASLDIWEKNVAVHRRGLELIRPGARCK 300
Db 241 VOSGDIIISLNTFPMIFGYTTLERTLFCDHVDASLDIWEKNVAVHRRGLELIRPGARCK 300
Qy 301 DIAELNEMRYEMDLKTRSPGYGSHFCVLCYYYGREGAVELREDIDITVLQPGMVVSWEP 360
Db 301 DIAELNEMRYEMDLKTRSPGYGSHFCVLCYYYGREGAVELREDIDITVLQPGMVVSWEP 360
Qy 361 MVMLEPGAPGAGYREHDLIVGEDGAMNTGPPGPEHNIRN 404
Db 361 MVMLEPGAPGAGYREHDLIVGEDGAMNTGPPGPEHNIRN 404

RESULT 9
; Sequence 21632, Application US/10369493
; Publication No. US2003023675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng

;; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
;; FILE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
;; FILE REFERENCE: 38-10(52052)B
;; CURRENT APPLICATION NUMBER: US/10/369,493
;; CURRENT FILING DATE: 2003-02-28
;; PRIOR APPLICATION NUMBER: US 60/360,039
;; PRIOR FILING DATE: 2002-02-21
;; NUMBER OF SEQ ID NOS: 47374
;; SEQ ID NO 21632
;; LENGTH: 351
;; TYPE: PRF
;; ORGANISM: Pyrococcus abyssi
US-10-369-493-21632

Query Match 11.3% Score 248.5; DB 15; Length 351;
Best Local Similarity 25.3%; Pred. No. 2e-15; Indels 41; Gaps 13;
Matches 92; Conservative 70; Mismatches 161;

QY 36 WMAKNNVDALFTSYHCINYYSGMLCYFGKRYGVNIDHNNATTISAGIDGQPMRRSFG 95
DB 14 FNNESIDAVLITKNPNYYLSGASPLAGY---ILVNDGATLVPELEVMABES-- 68
QY 96 DNTYTTDM-RDNFRAVRQLTTAKRIGIEPDHVNLPFRQLEALPGVEFVDSQSM 154
DB 69 -KIPVEKRRRDDFKVPE---GVKVGIG-GLSYSFVEDLKEKGISEKRVVDYIK 122
QY 155 WRTITKSLSEOKLIREGARVCDVGACAAIKAGVPEHVALATTNMIKIAKSPFV 214
DB 123 EMRVKSSSEETIKELKACEIDAKVMAAIEETBEKREITAKYELMRONGAEKPAFD 182
QY 215 ELMDTWTFQSGINTDGAHNPVTRIVQSGDILSINTEPMIFGYTALERTLFCDHVDA 274
DB 183 -----TIASGYNSALPHGVASDRIKSGDLVIDALVNHVMSDITRTVVGSPNEK 236
QY 275 SLDIWEKNAVHRGELIKRGARCKDIAELNMYREMDLKTSPFGYSGSPVLCY 334
DB 237 QREIYEIYLEAKQAVEAARPGITTK---ELDSIAR--NIIK--EYGYGYF--IHSL 285
QY 335 GREAGVELRE-----DIDTELKPGVVMSEPMVMLEPGMGAGGYEHDLIYGEDEGA 387
DB 286 GHGVGLEHHEMPGVSYQD-ETVLKRGWVITLPGIYIK---FGGVRIEDTLVITKTGA 340
QY 388 ENIT 391
DB 341 RLRT 344

RESULT 10
US-09-815-242-10645
;; Sequence 10645, Application US/09815242
;; Patent No. US20020061569A1
;; GENERAL INFORMATION:
;; APPLICANT: Haselbeck, Robert
;; APPLICANT: Ohlsen, Karl L.
;; APPLICANT: Zykkind, Judith W.
;; APPLICANT: Wall, Daniel
;; APPLICANT: Trawick, John D.
;; APPLICANT: Carr, Grant J.
;; APPLICANT: Yamamoto, Robert T.
;; APPLICANT: Xu, H. Howard
;; TITLE OF INVENTION: Identification of Essential Genes in
;; FILE REFERENCE: ELITRA 011A
;; CURRENT APPLICATION NUMBER: US/09/815,242
;; CURRENT FILING DATE: 2001-03-21
;; PRIOR APPLICATION NUMBER: 60/191,078
;; PRIOR FILING DATE: 2000-03-21
;; PRIOR APPLICATION NUMBER: 60/206,848
;; PRIOR FILING DATE: 2000-05-23
;; PRIOR APPLICATION NUMBER: 60/207,727
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: 60/242,578
;; PRIOR FILING DATE: 2000-10-23

;; PRIOR APPLICATION NUMBER: 60/253,625
;; PRIOR FILING DATE: 2000-11-27
;; PRIOR APPLICATION NUMBER: 60/257,931
;; PRIOR FILING DATE: 2000-12-22
;; PRIOR APPLICATION NUMBER: 60/269,308
;; PRIOR FILING DATE: 2001-02-16
;; NUMBER OF SEQ ID NOS: 14110
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 10645
;; LENGTH: 354
;; TYPE: PRF
;; ORGANISM: Enterococcus faecalis
US-09-815-242-10645

Query Match 10.5% Score 229.5; DB 9; Length 354;
Best Local Similarity 24.7%; Pred. No. 1.5e-13;
Matches 98; Conservative 59; Mismatches 164; Indels 75; Gaps 14;

QY 26 MTRQNDVRGMMAKNNVDALFTSYHCINYYSGMLCYFGKRYGVNIDHNNATTISAGID 85
DB 1 MAAVEKRLRELMKENNLQGLVTSPLYMLRYLTN-----FTGTGLAM-----IAMD 46
QY 86 GGGPMRRSFGDNITYTTDMRDNFRAVRQLT--TGAKRIGIEPDHV-----NIDF 133
DB 47 -----KAFITDFPYTE-----QAAEQATGFTIVKNTGIFPEVADLAERLQDNLAF 94
QY 134 RRQ-----LEALPGVEFVDISQSMWRTITKSLSEOKLIREGARVCDVGACAA 184
DB 95 EETQVSFADYSLEHILP-CELVPMGLIELEREYKDEEVAITIKACALDQGAFLVE 153
QY 185 AIKAGVEHVALATTNMIKIAKSPFVELMDTWTFQSGINTDGAHNPVTRIVQSG 244
DB 154 MIKGMTEIEVANQLDPMRSGKASGVSEFTIV-----ASGLNSAMPHGVASHKVIKRG 207
QY 245 DILSINTEPMIFGYTALERTLFCDHVDASLDIWEKNAVHRGELIKRGARCKDIAI 304
DB 208 ELITLDFRCYEGYGVSDMTRFPAISIQPKLKEIYDYLEQLVLAERKGL---TGI 263
QY 305 ELNMYREMDLKTSPFGYSGSPVLCYHGREAGVELRE-----IDTELKPGVVM 358
DB 264 QLDALAND---HIASGYGAFG---HSTHGIGLEIHEPNVNSFRADKQFVGPNVITD 316
QY 359 EPMVMLEPGMGAGGYEHDLIYGEDEGAENITGPP 394
DB 317 EPGIYL-----PGIGVRIEDDLSTIAGNRVLTHAP 348

RESULT 11
US-09-815-242-10235
;; Sequence 10235, Application US/09815242
;; Patent No. US20020061569A1
;; GENERAL INFORMATION:
;; APPLICANT: Haselbeck, Robert
;; APPLICANT: Ohlsen, Karl L.
;; APPLICANT: Zykkind, Judith W.
;; APPLICANT: Wall, Daniel
;; APPLICANT: Trawick, John D.
;; APPLICANT: Carr, Grant J.
;; APPLICANT: Yamamoto, Robert T.
;; APPLICANT: Xu, H. Howard
;; TITLE OF INVENTION: Identification of Essential Genes in
;; FILE REFERENCE: ELITRA 011A
;; CURRENT APPLICATION NUMBER: US/09/815,242
;; CURRENT FILING DATE: 2001-03-21
;; PRIOR APPLICATION NUMBER: 60/191,078
;; PRIOR FILING DATE: 2000-03-21
;; PRIOR APPLICATION NUMBER: 60/206,848
;; PRIOR FILING DATE: 2000-05-23
;; PRIOR APPLICATION NUMBER: 60/207,727
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: 60/242,578
;; PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10235
LENGTH: 361
TYPE: PRT
ORGANISM: Escherichia coli
US-09-815-242-10235

Query Match 10.5%; Score 229; DB 9; Length 361;
Best Local Similarity 23.9%; Pred. No. 1.7e-13;
Matches 93; Conservative 64; Mismatches 162; Indels 70; Gaps 13;

QY 33 VRGMANNVDAALFTSYH---CINYSGLYCYFGKRG-WIDHNNATTISAGIDG 87
DB 7 LRDLWKQQLDAVLISRKQKPHLGISTGSGYVVISRESAHILVDSRYVEVEARAQGY 66
QY 88 QPARRSPEDNTTYDWRDNFYRAVROLTTGAKRIGIEFDHVN-----LDFRQLE 139
DB 67 QHLHLDATNTLTTI---VNOIADDEQLQT---LGFEGQVSWETAHRQSELMAKLVS 118
QY 140 ALPGVEFYDISQPSMMRTIKSLBEOKLIRGARVCDVGAAACAARAGVPEHEVALAT 199
DB 119 ATPDV-----LRQITPEVEKIRLACGADRGAEHIRFIQAGSERIAEL 167
QY 200 TNAMIREAKSPFYELMDTWTWFOGSGINTGAGNPVTRIVQSDILSLTFPIRIFY 259
DB 168 EFWFRQGAERAKSEFDIV-----ASGWRGALPHGKADKIYAASEFTVLDLFGALYQCYC 221
QY 260 TALERTLECD-----HNDASLDIWEKNVAVHRRGLIKPGARKDIAELNENYR 311
DB 222 SDMTRTLIVNKGVSASESHL---LFPVQIVLQAGLAISAIRPVRCCQVDDAARVIT 278
QY 312 EMDLKYRSPCYGHSFGVLCYHVGREAGVELRED-----IDTELKPMVVMSEPMWLP 365
DB 279 E-----AGYGVFG---HNTGHAIGIEVHEDRFSPRDTTLQPGMLTVERGIYL- 326
QY 366 EGMFGAGYRHHDLIYGEDGAENITGFP 394
DB 327 ---PGGGGVRIEDVVLVTPGAGVLYAMP 352

RESULT 12

US-10-156-761-9699
Sequence 9699, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OMIURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 9699
LENGTH: 381
TYPE: PRT
ORGANISM: Streptomyces avermitilis
US-10-156-761-9699

Query Match 9.3%; Score 203.5; DB 14; Length 381;
Best Local Similarity 24.3%; Pred. No. 5.8e-11;
Matches 97; Conservative 57; Mismatches 186; Indels 59; Gaps 15;

QY 14 GEXDYPSPDAEMTRQNDVRGMANNVDAALFTSYHCINYSGLYCYFGKRGKGYV 72
DB 9 GTGDPAFTADYAAARRRRAAGSAAADGDLGVLAPEPDLVWLGVPEVERLTLLVLR 68
QY 73 -DHNNATTI-----SAGIDGQFPWRRSPEDNTTYDWR---RDNFYAVROLTTGAKRIG 123
DB 69 AGHDPVLVPTLEAPDAABAAGP-----ALTRDMTGDKDP-YEAAATILDRSGFSG 120
QY 124 IEPHNVLDPRRQLEELPGVEFYDISQPSMMRTIKSLBEOKLIRGARVCDVGAAACA 183
DB 121 IS-DNAMAMHLIGSGLPGTRYVALTEALPMLKAVDALE-----RIAAAGAAADA 173
QY 184 AAKR-----AGVEHEVAITTNAMIREIAKSPPELMDTWTWFOGSGINTDGAHPVT 237
DB 174 TFEIRKVPFRAGRETDVA-----ADLAELRQGHGQV---DFTVSGSPRGANPHHEAG 226
QY 238 NRIVQSDILSLTFPMIFGYTALERTLPCDHVDAASLDIWEKNVAVHRRGLELIRGA 297
DB 227 DVIERGDVVLDRGKLGKGYSDTSRTVHGEDEEEKVHDLVRAAQAEGFRAVRGV 286
QY 298 RCKDIAELNENYRHHLLKRSFGYHSGVLCYHVGREAGVELRED---IDTE---LK 351
DB 287 ACQVDRAARAVAD-----AGYDRP---IHRTHGIGVTTHEPWTIEBORALV 335
QY 352 PGWVSMPEMWLPBGMFGAGYRHHDLIYGEDGAENI 390
DB 336 PGMCFSPVPGVYL---PGRFGRVRIEDIVVTDDGGRL 370

RESULT 13

US-09-815-242-5287
Sequence 5287, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zykkind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELTRA, 011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5287
LENGTH: 350
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-5287

Query Match 9.1%; Score 200; DB 9; Length 350;
Best Local Similarity 21.7%; Pred. No. 1.1e-10; Indels 62; Gaps 12;
Matches 83; Conservative 70; Mismatches 168; Indels 62; Gaps 12;

QY 29 RONDVGWMAKNVDAALFTSYHCINYYSGMLCYCFGRKYGVNIDHNNATTISAGIDGQ 88
DB 1 RTQVHRIIEQKHDAIITLSDYNNRYLSG---FTGSGALIIISKDQYILIT----- 49
QY 89 PWRBSFGNITYTDM-----RDNFYRAVRQL--TTGAKRIGIEFDHYNLDR 134
DB 50 -----DFRYIDATKQAPNYEIIINRSTIIIGELKELHOENENWGFEGHWSYDXY 101
QY 135 RQLEBALPGVEFVDSQPSMMWRRTIKSLSEQGLIREGARVCVGAACAALAKAVPEHE 194
DB 102 LELNKS--RISLISISNTVDKIRVDKADDELALIQKANIYDETEYIITLVKAGMTEKE 159
QY 195 VAATTNMIREIAKSPFVELMDTWTFQSGINTDGAHNPVTNRIVOSGDIISLNTPEM 254
DB 160 LKALIESKML-ELGADGSPFD-----TIVASGHRGALPHGVASDKIEKGDWITLDFGAY 213
QY 255 IFGYTALERTLPCHVDASLDIWEKNVAHRRGLIELIKPGARCKDIAIEINENYREND 314
DB 214 YNGYCSDIITRFPAIGEPDPKKEIYQIVLESQMKAINERPGMTGAE-ADALSRNYLE-- 270
QY 315 LKXRSFGYGRSPGYLCHYGRGAEVBLE-----DIDTELKPGWVVSMEPMWLPBGM 368
DB 271 -----SKGYGKEFG---HSLHGIGLEIHEGPMALARTIQDKLVANNCTVBERGVYI----- 318
QY 369 PGAGYREHDILIVGEDGAENIT 391
DB 319 EGLGGIRIEDDILITENGCOVFT 341

Query Match 9.1%; Score 200; DB 9; Length 353;
Best Local Similarity 21.7%; Pred. No. 1.2e-10; Indels 62; Gaps 12;
Matches 83; Conservative 70; Mismatches 168; Indels 62; Gaps 12;

QY 29 RONDVGWMAKNVDAALFTSYHCINYYSGMLCYCFGRKYGVNIDHNNATTISAGIDGQ 88
DB 3 RTQVHRIIEQKHDAIITLSDYNNRYLSG---FTGSGALIIISKDQYILIT----- 51
QY 89 PWRBSFGNITYTDM-----RDNFYRAVRQL--TTGAKRIGIEFDHYNLDR 134
DB 52 -----DFRYIDATKQAPNYEIIINRSTIIIGELKELHOENENWGFEGHWSYDXY 103
QY 135 RQLEBALPGVEFVDSQPSMMWRRTIKSLSEQGLIREGARVCVGAACAALAKAVPEHE 194
DB 104 LELNKS--RISLISISNTVDKIRVDKADDELALIQKANIYDETEYIITLVKAGMTEKE 161
QY 195 VAATTNMIREIAKSPFVELMDTWTFQSGINTDGAHNPVTNRIVOSGDIISLNTPEM 254
DB 162 LKALIESKML-ELGADGSPFD-----TIVASGHRGALPHGVASDKIEKGDWITLDFGAY 215
QY 255 IFGYTALERTLPCHVDASLDIWEKNVAHRRGLIELIKPGARCKDIAIEINENYREND 314
DB 216 YNGYCSDIITRFPAIGEPDPKKEIYQIVLESQMKAINERPGMTGAE-ADALSRNYLE-- 272
QY 315 LKXRSFGYGRSPGYLCHYGRGAEVBLE-----DIDTELKPGWVVSMEPMWLPBGM 368
DB 273 -----SKGYGKEFG---HSLHGIGLEIHEGPMALARTIQDKLVANNCTVBERGVYI----- 320
QY 369 PGAGYREHDILIVGEDGAENIT 391
DB 321 EGLGGIRIEDDILITENGCOVFT 343

RESULT 14
US-09-815-242-12580
Sequence 12580, Application US/09815242
Patent No. US2002061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyekind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.01A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12580
LENGTH: 353
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-12580

RESULT 15
US-10-282-122A-70734
Sequence 70734, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl
APPLICANT: Zyekind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Foreyth, R.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.03A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 7, 2005, 19:36:24 ; Search time 5402 seconds
(without alignments)
10871.478 Million cell updates/sec

Title: US-10-807-228a-2

Perfect score: 1212

Sequence: 1 ATGACTGACGACATGTGCA.....AACACACATGATCCGCAAC 1212

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba: *
2: gb_hcg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1210.4	99.9	1212	6	AR100122
2	1210.4	99.9	1212	6	E12280
3	1210.4	99.9	1212	6	E13584
4	1210.4	99.9	1212	6	E16405
5	1210.4	99.9	1212	6	BD017699
6	1210.4	99.9	1212	6	BD105662
7	1208.8	99.7	1215	1	AB016788
8	1188	98.0	1215	6	E11155
9	984.8	81.3	1215	6	AX721933
10	984.8	81.3	1215	6	AX721949
11	981.6	81.0	1215	6	AX721941
12	980	80.9	1212	6	AX721943
13	980	80.9	1212	6	AX721945
14	980	80.9	1212	6	AX721947
15	978.4	80.7	1212	6	AX721953
16	976.8	80.6	1212	6	AX721951
17	643.4	53.1	1723	1	FVBCRI
18	643.4	53.1	1723	1	FVBCRE
19	643.4	53.1	1956	1	D00656 Flavobacter

20	638.8	52.7	1212	6	A10619
21	638.8	52.7	1212	6	E00904
22	638.8	52.7	1212	6	I01844
23	637.2	52.6	1212	6	A01506
24	582.6	48.1	300029	1	AE016787
25	572	47.2	3774	1	AF170566
26	570.4	47.1	1810	1	AF072304
27	566.2	46.7	1134	6	E01576
28	461	38.0	1233	6	E01828
29	461	38.0	1584	1	BACR2
30	376	31.0	8127	1	AB007122
31	374.4	30.9	1236	6	E17219
32	95.8	7.9	110000	1	AP006618
33	86	7.1	303855	1	AE017230
34	79.8	6.6	291000	1	SC0939105
35	77.4	6.4	302998	1	AE016821
36	76	6.3	321250	1	SC0939111
37	74.4	6.1	125020	9	AF429315
38	73.8	6.1	2783	8	AK119861
39	73.8	6.1	4740	8	AF271357
40	73.8	6.1	146841	8	AP003629
41	73.2	6.0	786	6	AX928131
42	73.2	6.0	786	6	AX928133
43	73.2	6.0	14537	1	AE004785
44	72.6	6.0	110000	1	RM5591985_04
45	72.4	6.0	14807	1	AB024601

ALIGNMENTS

RESULT 1	AR100122	1212 bp	DNA	linear	PAT 14-FEB-2001
LOCUS	AR100122	Sequence 2 from patent US 6080553.			
DEFINITION	AR100122				
ACCESSION	AR100122.1	GI:12810570			
VERSION					
KEYWORDS					
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	1 (bases 1 to 1212)				
AUTHORS	Sogabe,A., Hattori,T., Nishiya,Y. and Kawamura,Y.				
TITLE	Creatine amidinohydrolase, production thereof and use thereof				
JOURNAL	Patent: US 6080553-A 2 27-JUN-2000;				
FEATURES	Location/Qualifiers				
SOURCE	1..1212				
ORIGIN	/organism="unknown"				
	/mol_type="unassigned DNA"				
Query Match	99.9%; Score 1210.4; DB 6; Length 1212;				
Best Local Similarity	99.9%; Pred. No. 5.7e-139;				
Matches 1211; Conservative	0; Mismatches 1; Indels 0; Gaps 0;				
QY	1 ATGACTGACGACATGTGCAAGTGAAGGCAACAGCGGAGAAATTTATTCGCG 60				
DB	1 ATGACTGACGACATGTGCAAGTGAAGGCAACAGCGGAGAAATTTATTCGCG 60				
QY	61 TTTTCGATGCGGATGACCCCGGCCAAACAGCTTCGCGGTGATGGCCAAAGAC 120				
DB	61 TTTTCGATGCGGATGACCCCGGCCAAACAGCTTCGCGGTGATGGCCAAAGAC 120				
QY	121 AATGTCATGCGGCGCTGTTCACTTTATCATGTCATCACTATTCGCGGCTG 180				
DB	121 AATGTCATGCGGCGCTGTTCACTTTATCATGTCATCACTATTCGCGGCTG 180				
QY	181 TACGCTATTTCGACCGAGTACGATGATGATGATGATGATGATGATGATGAT 240				
DB	181 TACGCTATTTCGACCGAGTACGATGATGATGATGATGATGATGATGATGAT 240				
QY	241 TCGCGCGGATGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300				

QY 481 TCGCTGAAGAGCAGAGTGTATCCGCGAAGCGCCGCGTGTGTGACGTCCGCGCGCG 540
 DB 481 TCGCTGAAGAGCAGAGTGTATCCGCGAAGCGCCGCGTGTGTGACGTCCGCGCGCG 540
 QY 541 GCTTCGCGCGTGCATCAAGCGCGCGTGCCTCGAGCATGAAGTGGCGATCCACAC 600
 DB 541 GCTTCGCGCGTGCATCAAGCGCGCGTGCCTCGAGCATGAAGTGGCGATCCACAC 600
 QY 601 AATGCGATGATCCGCGAATCCGCAATCCGCTTCGTGTGAGCTGATGGAACCTCG 660
 DB 601 AATGCGATGATCCGCGAATCCGCAATCCGCTTCGTGTGAGCTGATGGAACCTCG 660
 QY 661 ACTGCTTCAGTCCGCGATCAACACGCGCGCGCAATCCGCTTCGTGTGAGCTGATGGAACCTCG 720
 DB 661 ACTGCTTCAGTCCGCGATCAACACGCGCGCGCAATCCGCTTCGTGTGAGCTGATGGAACCTCG 720
 QY 721 GTGCAATCCGCGGATCATCTTTCGCTCAACACCTTCCGATGATCTTCCGCTACTAC 780
 DB 721 GTGCAATCCGCGGATCATCTTTCGCTCAACACCTTCCGATGATCTTCCGCTACTAC 780
 QY 781 GCGCTGAGCGCGCATCTTTCGCGACCATGTGATGACCGCGCTGACATCTGGAG 840
 DB 781 GCGCTGAGCGCGCATCTTTCGCGACCATGTGATGACCGCGCTGACATCTGGAG 840
 QY 841 AAGAAGTGGCGCTGATCGCGCGCGCTGAGCTGATCAAGCGCGCGCGCTGCAAG 900
 DB 841 AAGAAGTGGCGCTGATCGCGCGCGCTGAGCTGATCAAGCGCGCGCGCTGCAAG 900
 QY 901 GACATCGGCATCGAGCTCAACGAGATGTACCGCGAGTGGAGCTGCTGAAGTACCGCTCC 960
 DB 901 GACATCGGCATCGAGCTCAACGAGATGTACCGCGAGTGGAGCTGCTGAAGTACCGCTCC 960
 QY 961 TTCCGCTATGCGCACTCTTTCGCGCGTGTGTGCGCACTGACGCTGCGAGCGCGCGCTG 1020
 DB 961 TTCCGCTATGCGCACTCTTTCGCGCGTGTGTGCGCACTGACGCTGCGAGCGCGCGCTG 1020
 QY 1021 GAGCTGCGCGAGCATCGACACCGAGCTGAAGCGCGCGCATGTGTCTCCATGAGCGCG 1080
 DB 1021 GAGCTGCGCGAGCATCGACACCGAGCTGAAGCGCGCGCATGTGTCTCCATGAGCGCG 1080
 QY 1081 ATGCTGATGCTGCGCGAGCGCATGCGCGCGTGCCTGATCGGAGCAGCATCTCTG 1140
 DB 1081 ATGCTGATGCTGCGCGAGCGCATGCGCGCGTGCCTGATCGGAGCAGCATCTCTG 1140
 QY 1141 ATGCTGCGGAGAGCATGCGCGCGTGCCTGATCGGAGCAGCATCTCTG 1200
 DB 1141 ATGCTGCGGAGAGCATGCGCGCGTGCCTGATCGGAGCAGCATCTCTG 1200
 QY 1201 ATCATCCGCAAC 1212
 DB 1201 ATCATCCGCAAC 1212

RESULT 3
 E13584
 LOCUS 1212 bp DNA linear PAT 27-APR-1998
 DEFINITION Alcaligenes faecalis gene for creatine amidinohydrolase.
 ACCESSION E13584
 VERSION E13584.1 GI:3252389
 KEYWORDS JP 1997215494-A/1.
 SOURCE Alkaligenes faecalis
 ORGANISM Alkaligenes faecalis
 Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 Alcaligenaceae; Alcaligenes.
 REFERENCES
 1 (bases 1 to 1212)
 Sugabe, A., Hattori, T., Nishiy, Y. and Kawamura, Y.
 NEW CREATINE AMIDINOHYDROLASE ITS PRODUCTION AND ITS USE
 JOURNAL Patent: JP 1997215494-A 1 19-AUG-1997;
 TOYOBO CO LTD
 COMMENT
 OS Alkaligenes faecalis
 PN JP 1997215494-A/1
 PD 19-AUG-1997
 PF 13-FEB-1996 JP 1996025435

PI SOGABE ATSUSHI, HATTORI TAKASHI, NISHIYA YOSHIKAKI, PI
 KAWAMURA YOSHIKAKI
 PC C12N9/78, C12N15/09, (C12N9/78, C12R1:05), (C12N15/09, C12R1:19);
 CC strandedness: Double;
 CC topology: linear;
 FH key Location/Qualifiers
 FT source 1..1212
 /strain="TE3581"
 /organism="Alcaligenes faecalis" FT
 /location/Qualifiers
 1..1212
 /organism="Alcaligenes faecalis"
 /mol_type="genomic DNA"
 /db_xref="taxon:511"

ORIGIN
 Query Match 99.9%; Score 1210.4; DB 6; Length 1212;
 Best Local Similarity 99.9%; Pred. No. 5.7e-139;
 Matches 1211; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ATGACTGACGACATGTTGACAGTGTATGAAATGAGCAACACGCGAGAAAGATTATTCGCG 60
 1 ATGACTGACGACATGTTGACAGTGTATGAAATGAGCAACACGCGAGAAAGATTATTCGCG 60
 61 TTTTCGAGTCCGAGATGACCGCGCGCAACAGCGTTCGCGGCTGAGTGGCCAGAAG 120
 61 TTTTCGAGTCCGAGATGACCGCGCGCAACAGCGTTCGCGGCTGAGTGGCCAGAAG 120
 121 AATGTGATGCGCGGCTGTTTACCTTATGATGATCACTGATCACTATTCGCGCTG 180
 121 AATGTGATGCGCGGCTGTTTACCTTATGATGATCACTGATCACTATTCGCGCTG 180
 121 AATGTGATGCGCGGCTGTTTACCTTATGATGATCACTGATCACTATTCGCGCTG 180
 181 TACTGCTATTTGCGACGCAATGATGAGTATGATGACCAACCAAGCCACGAGATT 240
 181 TACTGCTATTTGCGACGCAATGATGAGTATGATGACCAACCAAGCCACGAGATT 240
 181 TACTGCTATTTGCGACGCAATGATGAGTATGATGACCAACCAAGCCACGAGATT 240
 241 TGGCGCGCGCATGACGCGCGCGAGCGCTGCGCGCGCGCTTCCGCGCAACATCACTAC 300
 241 TGGCGCGCGCATGACGCGCGCGAGCGCTGCGCGCGCGCTTCCGCGCAACATCACTAC 300
 241 TGGCGCGCGCATGACGCGCGCGAGCGCTGCGCGCGCGCTTCCGCGCAACATCACTAC 300
 301 ACCGACTGCGCGCGCGCAATTTCTATGCGCGCGTGCCTGAGCTGACCGAGCGCGCAAG 360
 301 ACCGACTGCGCGCGCGCAATTTCTATGCGCGCGTGCCTGAGCTGACCGAGCGCGCAAG 360
 301 ACCGACTGCGCGCGCGCAATTTCTATGCGCGCGTGCCTGAGCTGACCGAGCGCGCAAG 360
 361 CGCATCGGCATGAGTGTGACACAGTCAATCTGCACTTCCGCGCAAGCTGAGGAGGCC 420
 361 CGCATCGGCATGAGTGTGACACAGTCAATCTGCACTTCCGCGCAAGCTGAGGAGGCC 420
 361 CGCATCGGCATGAGTGTGACACAGTCAATCTGCACTTCCGCGCAAGCTGAGGAGGCC 420
 421 CTACCGGCGCTGAGTTCGTGCAATGACGCGCGCTGAGTGTGAGTGTGAGTGTGAGTGTGAG 480
 421 CTACCGGCGCTGAGTTCGTGCAATGACGCGCGCTGAGTGTGAGTGTGAGTGTGAGTGTGAG 480
 421 CTACCGGCGCTGAGTTCGTGCAATGACGCGCGCTGAGTGTGAGTGTGAGTGTGAGTGTGAG 480
 481 TCGCTGAAGAGCAGAGTGTATCCGCGAAGCGCGCGCTGATGAGTGTGAGTGTGAGTGTGAG 540
 481 TCGCTGAAGAGCAGAGTGTATCCGCGAAGCGCGCGCTGATGAGTGTGAGTGTGAGTGTGAG 540
 481 TCGCTGAAGAGCAGAGTGTATCCGCGAAGCGCGCGCTGATGAGTGTGAGTGTGAGTGTGAG 540
 541 GCTTCGCGGCTGCTCATGAGCGCGCGTGCCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 600
 541 GCTTCGCGGCTGCTCATGAGCGCGCGTGCCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 600
 541 GCTTCGCGGCTGCTCATGAGCGCGCGTGCCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 600
 601 AATGCGATGATCCGCGAATCCGCAATCCGCTTCGTGTGAGCTGATGGAACCTCG 660
 601 AATGCGATGATCCGCGAATCCGCAATCCGCTTCGTGTGAGCTGATGGAACCTCG 660
 601 AATGCGATGATCCGCGAATCCGCAATCCGCTTCGTGTGAGCTGATGGAACCTCG 660
 661 ACTGCTTCAGTCCGCGATCAACACGCGCGCGCAATCCGCTTCGTGTGAGCTGATGGAACCTCG 720
 661 ACTGCTTCAGTCCGCGATCAACACGCGCGCGCAATCCGCTTCGTGTGAGCTGATGGAACCTCG 720
 661 ACTGCTTCAGTCCGCGATCAACACGCGCGCGCAATCCGCTTCGTGTGAGCTGATGGAACCTCG 720
 721 GTGCAATCCGCGGATCATCTTTCGCTCAACACCTTCCGATGATCTTCCGCTACTAC 780
 721 GTGCAATCCGCGGATCATCTTTCGCTCAACACCTTCCGATGATCTTCCGCTACTAC 780
 721 GTGCAATCCGCGGATCATCTTTCGCTCAACACCTTCCGATGATCTTCCGCTACTAC 780

Qy	781	GCGCTGAGAGCGCAACGCTGTTCTTGCGGACGATGTCGATGACGCGACGCTCGACATCTGGAG	840
Ds	781	GCGCTGAGAGCGGACCGCTGTTCTTGCGGACGATGTCGATGACGCGACGCTCGACATCTGGAG	840
Qy	841	AAGAACGTGAGCGGTGECATCGCGCGCGGCGTGGAGTCAAGCCGGGCGCGCGTGGCAAG	900
Ds	841	AAGAACGTGAGCGGTGECATCGCGCGGCGGTGGAGTCAAGCCGGGCGCGCGTGGCAAG	900
Qy	901	GACATCGGCATTCGAGCTCAACGAGATGTAAACCGGAGATGGAGACCTGCTGTGAATTAACGGCTCC	960
Ds	901	GACATCGGCATTCGAGCTCAACGAGATGTAAACCGGAGATGGAGACCTGCTGTGAATTAACGGCTCC	960
Qy	961	TTCCGCGTATGAGCACCTCTTCGCGCGCTGTGTGTGCACATCACTACCGGTGCGGAGGCGGCGTGTG	1020
Ds	961	TTCCGCGTATGAGCACCTCTTCGCGCGCTGTGTGTGCACATCACTACCGGTGCGGAGGCGGCGTGTG	1020
Qy	1021	GAGCTGCGCGAGAGACATCGACACCGAGCTGAAAGCCCGGACATGTGTGTTCATATGAGACCG	1080
Ds	1021	GAGCTGCGCGAGAGACATCGACACCGAGCTGAAAGCCCGGACATGTGTGTTCATATGAGACCG	1080
Qy	1081	ATGTGATGATGCTGCGCGGAGGAGCATGCGCGGTGCGCGGCGGCTATTCGAGACACGACATCTGTG	1140
Ds	1081	ATGTGATGATGCTGCGCGGAGGAGCATGCGCGGTGCGCGGCGGCTATTCGAGACACGACATCTGTG	1140
Qy	1141	ATCGTCGGGAGAGACGATGTCGCGAGAACATCAACCGGCTTCCGTTGCGGTCCGGAACACAAAC	1200
Ds	1141	ATCGTCGGGAGAGACGATGTCGCGAGAACATCAACCGGCTTCCGTTGCGGTCCGGAACACAAAC	1200
Qy	1201	ATCATCCGCAAC 1212	
Ds	1201	ATCATCCGCAAC 1212	

RESULT 4	E16405	LOCUS	1212 bp	DNA	linear	PAT 28-JUL-1999
DEFINITION	Alcallegenes faecalis gene for creatine amidinohydrolase.					

ACCESSION	E16405
VERSION	E16405.1
KEYWORDS	GI:5711088
SOURCE	JP 1998174585-A/1.
ORGANISM	Alcaligenes faecalis
	Alcaligenes faecalis
	Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
	Alcaligenaceae; Alcaligenes.
REFERENCE	1 (Babes 1 to 1212)
AUTHORS	Sogabe,A.; Nishiya,Y. and Kawamura,Y.
TITLE	STABLE CREATININE AMIDINOHYDROLASE
JOURNAL	Patent: JP 1998174585-A 1 30-JUN-1998;
COMMENT	TOYOBO CO LTD
	OS Alcaligenes faecalis

PN	JP	1998174585-A/1	
PD	30-JUN-1998		
PF	17-DEC-1996	JP	1996337027
PI	SOGABE ATSUSHI, NISHIYA YOSHIAKI, KAWAMURA YOSHIIISA PC		
CI2N9/78, C07H21/04, CI2N1/21, CI2N15/09//CI201/34, (CI2N9/78, PC			
CI2R1:19)			
PC	(CI2N1/21, CI2R1:19);		
CC	strandedness: Double;		
CC	topology: Linear;		
PH	Key	Location/Qualifiers	
PH			
FT	source	1. .1212	
FT	/strain='TE3581'	/organism='Alcaligenes faecalis'	FT
FT	mat_peptide	1. .1212	
FT	/product='creatinine amidinohydrolase'.		

FEATURES	Location/Qualifiers
source	1..1212
	/organism="Alcaligenes faecalis"
	/mol_type="genomic DNA"
	/db_xref="taxon:511"

ORIGIN

Query Match	99.9%	Score 1210.4;	DB 6;	Length 1212;
Best Local Similarity	99.9%;	Pred. No. 5.7e-139;		
Matches 1211; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

QY	1	ATGACTGACGACATGTTGACAGTGATGAAATGGCAACAGCGAGAAAGATTATTCGCGG	60
Db	1	ATGACTGACGACATGTTGACAGTGATGAAATGGCAACAGCGAGAAAGATTATTCGCGG	60
QY	61	TTTTTCGATGCGCGAGATGACCCCGCGCAAAACGATTCGGCGGTGGATGGCCAAAGAC	120
Db	61	TTTTTCGATGCGCGAGATGACCCCGCGCAAAACGATTCGGCGGTGGATGGCCAAAGAC	120
QY	121	AATGTCGATGGGAGCGCTGTTCACCTCTTATCACTGCATCAACTACTATTCGCGCTGCTG	180
Db	121	AATGTCGATGGGAGCGCTGTTCACCTCTTATCACTGCATCAACTACTATTCGCGCTGCTG	180
QY	181	TACTGCTATTTTCGAGCGCAAGTACGGCATGTGTATTCGACCAACCAACAGCGACGATT	240
Db	181	TACTGCTATTTTCGAGCGCAAGTACGGCATGTGTATTCGACCAACCAACAGCGACGATT	240
QY	241	TCGGCCGGGCATCGACGGCGGCGCAGCCCTGGGCGCGCAGCTTCGGCGCAACATCACTTAC	300
Db	241	TCGGCCGGGCATCGACGGCGGCGCAGCCCTGGGCGCGCAGCTTCGGCGCAACATCACTTAC	300
QY	301	ACCGACTGGCGCGCGCACAATTTCTATTCGGCGCGTGGCGCACTGACCAAGGCGGCCAAG	360
Db	301	ACCGACTGGCGCGCGCGCACAATTTCTATTCGGCGCGTGGCGCACTGACCAAGGCGGCCAAG	360
QY	361	CGCATTCGGGCATTCGAGTTTCGACACAGTCATCTTCGACCTTCGGCGCGCAGCTTCGAGAAAGCC	420
Db	361	CGCATTCGGGCATTCGAGTTTCGACACAGTCATCTTCGACCTTCGGCGCGCAGCTTCGAGAAAGCC	420
QY	421	CTACCGGAGCGCTGAGTTTCGTGCACATGACGCAAGCCCTCGATGTGTGATGGCGACCATCAAG	480
Db	421	CTACCGGAGCGCTGAGTTTCGTGCACATGACGCAAGCCCTCGATGTGTGATGGCGACCATCAAG	480
QY	481	TCGCTCCGAAGGCAAGAGCTGATTCGGGGAAGGCGCCCGCGTGTGACGTTCGGGCGGCGG	540
Db	481	TCGCTCCGAAGGCAAGAGCTGATTCGGGGAAGGCGCCCGCGTGTGACGTTCGGGCGGCGG	540
QY	541	GCCTGCGCGGCTGCATCAAGGCGGCGGTGCCGAGCATGAAATGGCGCATGCCACCAAC	600
Db	541	GCCTGCGCGGCTGCATCAAGGCGGCGGTGCCGAGCATGAAATGGCGCATGCCACCAAC	600
QY	601	AATCGCATGATCCGCGAGATGCCAAATTCGTTCCCTTCGTGAGCTGATGGAACCTCTGG	660
Db	601	AATCGCATGATCCGCGAGATGCCAAATTCGTTCCCTTCGTGAGCTGATGGAACCTCTGG	660
QY	661	ACCTGATTCCAGTGGGGCATCAACACGCGAGCGGCAATTCGGGTCAACACCGCATC	720
Db	661	ACCTGATTCCAGTGGGGCATCAACACGCGAGCGGCAATTCGGGTCAACACCGCATC	720
QY	721	GTGCAATCCGCGCAATCTCTTCGCTCAACACTTCCCGATGATCTTCGGGTACTACAC	780
Db	721	GTGCAATCCGCGCAATCTCTTCGCTCAACACTTCCCGATGATCTTCGGGTACTACAC	780
QY	781	GCGCTGAGCGGCAAGCTGTTCTGCGACATGTGATGACGCAAGCTTCGACATCTGGAG	840
Db	781	GCGCTGAGCGGCAAGCTGTTCTGCGACATGTGATGACGCAAGCTTCGACATCTGGAG	840
QY	841	AAGAACGTGGCCGTGACATCGCCCGGAGCTCGAGCTGATCAAGCCGGGCGCGCTGCAAG	900
Db	841	AAGAACGTGGCCGTGACATCGCCCGGAGCTCGAGCTGATCAAGCCGGGCGCGCTGCAAG	900
QY	901	GACATTCGCGCATCGAGCTCAACGAGATGTATCCGCGAGTGGGACCTGCTGAAATACCGCTCC	960
Db	901	GACATTCGCGCATCGAGCTCAACGAGATGTATCCGCGAGTGGGACCTGCTGAAATACCGCTCC	960
QY	961	TTCCGCTATGGGCACTCCTTCGCGGTCTGTGCACTACTACAGGTTCGCGAGGCGCGCTG	1020
Db	961	TTCCGCTATGGGCACTCCTTCGCGGTCTGTGCACTACTACAGGTTCGCGAGGCGCGCTG	1020
QY	1021	GAGCTGCGCGAGGACATTCGACACCGAGCTGAAGCCCGGACATGATGTGTCTCATGGAAGCG	1080

Db 1021 GAGCTGCGGAGACATCCACCGAGCTGAAGCCCGGATGGTGTCTCATGAGCCG 1080
Qy 1081 ATGTGATCTCTCCGAGGAGCATGCGCGTGCCTATCCGAGAGACATCTCG 1140
Db 1081 ATGTGATCTCTCCGAGGAGCATGCGCGTGCCTATCCGAGAGACATCTCG 1140
Qy 1141 ATGTGCGGAGAGACGCTCCGAGACATCACCGGCTTCCTCGTCCGTCGGAACAAC 1200
Db 1141 ATGTGCGGAGAGACGCTCCGAGACATCACCGGCTTCCTCGTCCGTCGGAACAAC 1200
Qy 1201 ATCATCCGCAAC 1212
Db 1201 ATCATCCGCAAC 1212

RESULT 5
BD017699
LOCUS BD017699 1212 bp DNA linear PAT 27-AUG-2002
DEFINITION Gene encoding creatine amidinohydrolase.
ACCESSION BD017699.1 GI:22558875
VERSION BD017699.1
KEYWORDS JP 2001252088-A/1.
SOURCE unclassified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 1212)
AUTHORS Sogabe,A., Yamamoto,K. and Kawamura,Y.
TITLE Gene encoding creatine amidinohydrolase
JOURNAL Patent: JP 2001252088-A 1 18-SEP-2001;
TOYOBO CO LTD
OS Alkaligenes faecalis TE3581 (FERM P14237)
PN JP 2001252088-A/1
PD 18-SEP-2001
PF 26-FEB-2001 JP 2001051054
PI ATSUSHI SOGABE, KAZUMI YAMAMOTO, YOSHIIHISA KAWAMURA PC
CI2N15/09, CI2N1/21, CI2N9/78/(CI2N15/09, CI2R1:425), (CI2N1/21, PC
CI2R1:425),
PC (CI2R9/78, CI2R1:05), CI2N15/00, (CI2N15/00, CI2R1:425) CC Gene
encoding creatine amidinohydrolase
FH Key Location/Qualifiers
FT source 1..1212
FT Location/Qualifiers
FT 1..1212
/organism="Alkaligenes faecalis TE3581 (FERM
P14237)"
/organism="unclassified"
/mol_type="Genomic DNA"
/db_xref="taxon:32644"

FEATURES
source
ORIGIN
Query Match 99.9%; Score 1210.4; DB 6; Length 1212;
Best Local Similarity 99.9%; Pred. No. 5,7e-139;
Matches 1211, Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 241 TCGGCGGATCGACGGCGGCGAGCCCTGCGCCGACGCTTCGGGACAAACATCACTAC 300
Qy 301 ACCGACTGGCGCGGAGCAATTTCTATCGCGCGCGTGGCGACGTGACACCGGGCGCAAG 360
Db 301 ACCGACTGGCGCGGAGCAATTTCTATCGCGCGCGTGGCGACGTGACACCGGGCGCAAG 360
Qy 361 CGCATCGGCATCGAGTTGACACCGTCAATCTGCACTTCCGCGCAGCTTCGAGAGGCC 420
Db 361 CGCATCGGCATCGAGTTGACACCGTCAATCTGCACTTCCGCGCAGCTTCGAGAGAGCC 420
Qy 421 CTACCGGCGCTCGAGTTGTCGACATCAGCCAGCCCTTGATGTGATGCGACATCAAG 480
Db 421 CTACCGGCGCTCGAGTTGTCGACATCAGCCAGCCCTTGATGTGATGCGACATCAAG 480
Qy 481 TCGCTGAAAGACAGACCTGATCCGGAGAGCGCGCGGTGTGAGTGTGGCGGCGCG 540
Db 481 TCGCTGAAAGACAGACCTGATCCGGAGAGCGCGCGGTGTGAGTGTGGCGGCGCG 540
Qy 541 GCGTCGCGGCTGCCATCAAGCGCGCGTGCCTCCGAGATGAAGTGGCATCGCACACC 600
Db 541 GCGTCGCGGCTGCCATCAAGCGCGCGTGCCTCCGAGATGAAGTGGCATCGCACACC 600
Qy 601 AATGCGATGATCCGCGAGATCGCAATCTGTTCCCTTCTGTGAGACTGATGACACTCG 660
Db 601 AATGCGATGATCCGCGAGATCGCAATCTGTTCCCTTCTGTGAGACTGATGACACTCG 660
Qy 661 ACGTGTTCAGTGGGCGATCAACCGAGCGCGGCAATCCGCTGACCAACCGCATC 720
Db 661 ACGTGTTCAGTGGGCGATCAACCGAGCGCGGCAATCCGCTGACCAACCGCATC 720
Qy 721 GTGCAATCCGCGCATCTTCTGCTCAACACTTCCGATGATCTTCGGCTACTACAC 780
Db 721 GTGCAATCCGCGCATCTTCTGCTCAACACTTCCGATGATCTTCGGCTACTACAC 780
Qy 781 GCGCTGAGCGCAGCTGTTCTGCGACCATGTGATGACGCGACCTTGACATCTGGAG 840
Db 781 GCGCTGAGCGCAGCTGTTCTGCGACCATGTGATGACGCGACCTTGACATCTGGAG 840
Qy 841 AAGAACGTCGCGCATCCGCGCGGCTGAGCGTCAAGCGCGGCGCGCGCTGCAAG 900
Db 841 AAGAACGTCGCGCATCCGCGCGGCTGAGCGTCAAGCGCGGCGCGCGCTGCAAG 900
Qy 901 GACATCGCATCGAGCTCAACGAGATGTACCGGAGTGGACCTGCTAAGTACCGCTCC 960
Db 901 GACATCGCATCGAGCTCAACGAGATGTACCGGAGTGGACCTGCTAAGTACCGCTCC 960
Qy 961 TTCGGCTATGGCACTCTTCTGCGGCTGTGCACTAATCGGTCGGAGCGCGCTG 1020
Db 961 TTCGGCTATGGCACTCTTCTGCGGCTGTGCACTAATCGGTCGGAGCGCGCTG 1020
Qy 1021 GAGCTGGCGGAGCATCGACATCGAGCTGAAGCCCGGATGTGTCTCATGAGCCG 1080
Db 1021 GAGCTGGCGGAGCATCGACATCGAGCTGAAGCCCGGATGTGTCTCATGAGAGCCG 1080
Qy 1081 ATGTGATGCTGCGGAGAGGATGCCGATGCCGCGGCTATCCGAGACATCTCTG 1140
Db 1081 ATGTGATGCTGCGGAGAGGATGCCGATGCCGCGGCTATCCGAGACATCTCTG 1140
Qy 1141 ATGTGCGGAGAGACGCTGCGAGACATCACCGGCTTCCTCGTCCGTCGGAACAAC 1200
Db 1141 ATGTGCGGAGAGACGCTGCGAGACATCACCGGCTTCCTCGTCCGTCGGAACAAC 1200
Qy 1201 ATCATCCGCAAC 1212
Db 1201 ATCATCCGCAAC 1212

RESULT 6
BD105662
LOCUS BD105662 1212 bp DNA linear PAT 27-AUG-2002
DEFINITION Stable creatine amidinohydrolase.
ACCESSION BD105662
VERSION BD105662.1 GI:22651236

KEYWORDS JP 2001346594-A/1.
SOURCE unidentified
ORGANISM unclassified.

REFERENCE 1 (bases 1 to 1212)
AUTHORS Sogabe, A., Nishihara, Y., and Kawamura, Y.
TITLE Stable creatine amidinohydrolase
JOURNAL Patent: JP 2001346594-A 1 18-DEC-2001;
TOYOBO CO LTD
OS Alcaligenes faecalis
PN JP 2001346594-A/1
PD 18-DEC-2001
PF 19-APR-2001 JP 2001121708
PI ATSUSHI SOGABE, YOSHIAKI NISHIYAMA, YOSHIOHISA KAWAMURA, PC
C12N15/09, C12N1/15, C12N1/21, C12N5/10, C12N9/80// PC
(C12N9/80, C12N1/05, C12N15/00, C12N5/00)
CC Strandedness: Double;
CC Topology: Linear;
CC Stable creatine amidinohydrolase
FH Key Location/Qualifiers
FT source 1..1212 /organism='Alcaligenes faecalis',
location/Qualifiers
1..1212 /organism='unidentified',
/mol_type='genomic DNA',
/db_xref='taxon:32644'

ORIGIN

Query Match 99.9%; Score 1210.4; DB 6; Length 1212;
Best Local Similarity 99.9%; Pred. No. 5.7e-139;
Matches 1211; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGACTGACGACATGTTGCACTGATGAATGGACAAACGCGAGAAAATTATTCGCG 60
DB 1 ATGACTGACGACATGTTGCACTGATGAATGGACAAACGCGAGAAAATTATTCGCG 60
QY 61 TTTTGGATGCGGAGATGACCCGCGCCAAACGACGTTCCGGGCTGATGAGCCCAAG 120
DB 61 TTTTGGATGCGGAGATGACCCGCGCCAAACGACGTTCCGGGCTGATGAGCCCAAG 120
QY 121 AATGCGATGCGGCGCTGTTCACTTATCACTGATCACTAATTAATTCGGGCTG 180
DB 121 AATGCGATGCGGCGCTGTTCACTTATCACTGATCACTAATTAATTCGGGCTG 180
QY 181 TACTGCTATTTGACGACGACGACGATGATGACCAACCAACGCGACGAT 240
DB 181 TACTGCTATTTGACGACGACGACGATGATGACCAACCAACGCGACGAT 240
QY 241 TCGCGCGGATGACGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
DB 241 TCGCGCGGATGACGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
QY 301 ACCGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
DB 301 ACCGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
QY 361 CGCATGCGGATGACGATTCGACGATTCGACGATTCGACGATTCGACGATTCG 420
DB 361 CGCATGCGGATGACGATTCGACGATTCGACGATTCGACGATTCGACGATTCG 420
QY 421 CTACCGGCGGCTGAGTTGCTGACATCAAGCGGCGGCGGCGGCGGCGGCGG 480
DB 421 CTACCGGCGGCTGAGTTGCTGACATCAAGCGGCGGCGGCGGCGGCGGCGG 480
QY 481 TCGCTGAGAGCAGACGATGATTCGCGGAGGCGGCGGCGGCGGCGGCGGCGG 540
DB 481 TCGCTGAGAGCAGACGATGATTCGCGGAGGCGGCGGCGGCGGCGGCGGCGG 540
QY 541 GCGTGGCGGCTGCGATCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600
DB 541 GCGTGGCGGCTGCGATCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600

QY 601 AATGCGATGATCCGCGAGATGCGCAATGTTCCCTTCGTGAGCTGATGACACTG 660
DB 601 AATGCGATGATCCGCGAGATGCGCAATGTTCCCTTCGTGAGCTGATGACACTG 660
QY 661 ACTGTTTCCAGTGGGATCAACACGAGCGGCGGCGGCAATTCGGTACCAACGCG 720
DB 661 ACTGTTTCCAGTGGGATCAACACGAGCGGCGGCGGCAATTCGGTACCAACGCG 720
QY 721 GTGCAATCCGCGGATCCCTTCGTCAACACCTTCGGAATGATCTTCGGTACTAC 780
DB 721 GTGCAATCCGCGGATCCCTTCGTCAACACCTTCGGAATGATCTTCGGTACTAC 780
QY 781 GCGCTGAGCGCAGCTGTTCTGCAACATGTCAGTACGCGACCTGACATCTGGAG 840
DB 781 GCGCTGAGCGCAGCTGTTCTGCAACATGTCAGTACGCGACCTGACATCTGGAG 840
QY 841 AAGAACGTGCGGCTGATTCGCGGCGGCGGCTGAGCTGATCAAGCGGCGGCGG 900
DB 841 AAGAACGTGCGGCTGATTCGCGGCGGCGGCTGAGCTGATCAAGCGGCGGCGG 900
QY 901 GACATCCGCGATGAGCTCAACGAGATGTAACCGGAGTGGGACCTGTAAGTACCG 960
DB 901 GACATCCGCGATGAGCTCAACGAGATGTAACCGGAGTGGGACCTGTAAGTACCG 960
QY 961 TTGCGCTATGCGCACTCCTTCGCGGCTGCTGTCCTGCACTACGATCGGAGCG 1020
DB 961 TTGCGCTATGCGCACTCCTTCGCGGCTGCTGTCCTGCACTACGATCGGAGCG 1020
QY 1021 GAGCTGCGGAGACATGACACCGAGCTGAAGCCCGGCGGCTATGCGAGACATCT 1080
DB 1021 GAGCTGCGGAGACATGACACCGAGCTGAAGCCCGGCGGCTATGCGAGACATCT 1080
QY 1081 ATGCTGATGCTGCGGAGGAGGATGCTCCGCGGCGGCGGCTATGCGAGACATCT 1140
DB 1081 ATGCTGATGCTGCGGAGGAGGATGCTCCGCGGCGGCGGCTATGCGAGACATCT 1140
QY 1141 ATGCTGCGGAGAGAGCGGTGCGGAGACATCAACCGGCTTCGCTGCTGCGAA 1200
DB 1141 ATGCTGCGGAGAGAGCGGTGCGGAGACATCAACCGGCTTCGCTGCTGCGAA 1200
QY 1201 ATCATCCGCAAC 1212
DB 1201 ATCATCCGCAAC 1212

RESULT 7
AB016788 1215 bp DNA linear BCT 06-JAN-2000
LOCUS Alcaligenes sp. gene for creatine amidinohydrolase, complete cds,
strain:KS-85.
DEFINITION
AB016788
ACCESSION
AB016788.1 GI:6681665
VERSION
AB016788.1
KEYWORDS
creatine amidinohydrolase.
SOURCE
Alcaligenes sp.
ORGANISM
Alcaligenes sp.
Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
Alcaligenaceae; Alcaligenes.
REFERENCE
1 (sites)
AUTHORS Furukawa, K., Ichikawa, T., Koyama, Y., and Suzuki, M.
TITLE Molecular cloning and sequence analysis of the gene encoding
thermostable creatinase from Alcaligenes sp. KS-85
JOURNAL Unpublished
2 (bases 1 to 1215)
REFERENCE
AUTHORS Koyama, Y. and Furukawa, K.
TITLE Direct Submission
JOURNAL Submitted (05-AUG-1998) Yasuji Koyama, Kikkoman Corporation,
Research and Development, Noda 399, Noda, Chiba 278-0037, Japan
(E-mail: dais-kin@92.so-net.ne.jp, Tel: +81-471-23-5571,
Fax: +81-471-23-5959)

FEATURES
source
1..1215
/organism='Alcaligenes sp.'
/mol_type='genomic DNA'

CDS

```
/strains="KS-85"  
/db_xref="taxon:512"  
1..1215  
/codon_start=1  
/transl_table=11  
/product="creatine amidinohydrolase"  
/protein_id="BA88830.1"  
/db_xref="GI:668166"  
translation="MTDDMLVMTKMHNEKDYSPSPDAEMTRBNDVYRGMXKNVDA  
ALFTSYHNTYNSGMLYRGRKGMVTDHNNATTISGIDGGQWRSPDNITYTD  
WRDNFRAVRQLTTGAKRIGIEPDVNLDRRLQELPGEFVDISQSPMKETIK  
SLEBQKLRGARVCDVGAACAIAKGVBEHVALPTNAMIRELKSPPELMD  
TWTFOSGINTDGMNPTNRIVOSGDLISNTPEMTIGYTALETTLFCDHVDLMD  
DIERKNVAVHRRGLEIKRGARCKDIAELNEMTREMOLLKTRSPGYRSLCHY  
GREGVELREDIDTELKRGVMVSMPEPMMLBGMGAGVREHDLIVGEGANITG  
PPGPEHNITRN"
```

ORIGIN

```
Query Match      99.7%; Score 1208.8; DB 1; Length 1215;  
Best Local Similarity 99.8%; Pred. No. 8.9e-139;  
Matches 1210; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 ATGACTGACGACATTTGACGATGTAATGACCAACGGGAGAAATATTTCGCG 60  
DB 1 ATGACTGACGACATTTGACGATGTAATGACCAACGGGAGAAATATTTCGCG 60  
QY 61 TTTTCGGATGCGGATGACCCCGCCGCAACGATTCGCGGTGATGGCCAGAAC 120  
DB 61 TTTTCGGATGCGGATGACCCCGCCGCAACGATTCGCGGTGATGGCCAGAAC 120  
QY 121 AATGTCGATGCGGCGCTGTTCACTTATGCACTGATCACTAATATTTCGCGCTG 180  
DB 121 AATGTCGATGCGGCGCTGTTCACTTATGCACTGATCACTAATATTTCGCGCTG 180  
QY 181 TACTGCTATTTTCGACGACGATGCGGATGATGACCAACGACGCAAGATTT 240  
DB 181 TACTGCTATTTTCGACGACGATGCGGATGATGACCAACGACGCAAGATTT 240  
QY 241 TCGGCGCGGATGACGCGGCGGACGCGCTGCGGCGGACCTTCGGGACAAATCACTAC 300  
DB 241 TCGGCGCGGATGACGCGGCGGACGCGCTGCGGCGGACCTTCGGGACAAATCACTAC 300  
QY 301 ACCGACTGCGCGCGGACCAATTTCTATGCGCGCGGCGGACGCTGACGACGCGGCG 360  
DB 301 ACCGACTGCGCGCGGACCAATTTCTATGCGCGCGGCGGACGCTGACGACGCGGCG 360  
QY 361 CGCATCGGCGATGAGTTGACCACTGATCTGACCTTCGCGCGGCGGACGCTGAGGAA 420  
DB 361 CGCATCGGCGATGAGTTGACCACTGATCTGACCTTCGCGCGGCGGACGCTGAGGAA 420  
QY 421 CTACCGGCGGCTGAGTTGCTGACATGACGACGCTTGTGATGATGATGATGATGATG 480  
DB 421 CTACCGGCGGCTGAGTTGCTGACATGACGACGCTTGTGATGATGATGATGATGATG 480  
QY 481 TCGCTCGAAGAGAGAAAGCTGATCGCGGAGCGCGCGGCTGTGTGAGTGTGAGCGCG 540  
DB 481 TCGCTCGAAGAGAGAAAGCTGATCGCGGAGCGCGCGGCTGTGTGAGTGTGAGCGCG 540  
QY 541 GCCTGCGGCGCTGCGATCAAGCGCGCGCTGCGCGGATGAGTGGCGATGCGCAAC 600  
DB 541 GCCTGCGGCGCTGCGATCAAGCGCGCGCTGCGCGGATGAGTGGCGATGCGCAAC 600  
QY 601 AATGCGATGATCGCGGAGATGCGCAATGCTTCCCTTGTGAGCTGATGATGATGATG 660  
DB 601 AATGCGATGATCGCGGAGATGCGCAATGCTTCCCTTGTGAGCTGATGATGATGATG 660  
QY 661 ACCTGATGATCGCGGAGATGCGCAATGCTTCCCTTGTGAGCTGATGATGATGATG 720  
DB 661 ACCTGATGATCGCGGAGATGCGCAATGCTTCCCTTGTGAGCTGATGATGATGATG 720  
QY 721 GTGCAATCGCGGACATCTTTCGCTCAACACTTCCGATGATCTTTCGCTCAACACC 780  
DB 721 GTGCAATCGCGGACATCTTTCGCTCAACACTTCCGATGATCTTTCGCTCAACACC 780
```

```
QY 781 GCGCTGAGCGGACGCTGTTCTGCGACCATGTCATGACGCGGACGCTGACATCTGGAG 840  
DB 781 GCGCTGAGCGGACGCTGTTCTGCGACCATGTCATGACGCGGACGCTGACATCTGGAG 840  
QY 841 AAGAACGTCGCGGTCATGCGCGGCGGATGATGATGATGATGATGATGATGATGATG 900  
DB 841 AAGAACGTCGCGGTCATGCGCGGCGGATGATGATGATGATGATGATGATGATGATG 900  
QY 901 GACATCGCATGCGATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 960  
DB 901 GACATCGCATGCGATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 960  
QY 961 TTCGCTATGCGCACTCTTTCGCGGCTGTCGCACTACTACGTCGCGGACGCGGCTG 1020  
DB 961 TTCGCTATGCGCACTCTTTCGCGGCTGTCGCACTACTACGTCGCGGACGCGGCTG 1020  
QY 1021 GAGCTGCGGAGGACATGACATGACATGACATGACATGACATGACATGACATGACATG 1080  
DB 1021 GAGCTGCGGAGGACATGACATGACATGACATGACATGACATGACATGACATGACATG 1080  
QY 1081 ATGCTGATGCTGCGGAGGCGATGCGCGGCGGCTATCGGAGCAAGACATCGCTG 1140  
DB 1081 ATGCTGATGCTGCGGAGGCGATGCGCGGCGGCTATCGGAGCAAGACATCGCTG 1140  
QY 1141 ATCGTGGGAGGAGCGGTCGGAACATGACCGGCTTCCGTTGCTCCGGAACACAC 1200  
DB 1141 ATCGTGGGAGGAGCGGTCGGAACATGACCGGCTTCCGTTGCTCCGGAACACAC 1200  
QY 1201 ATCATCCGCAAC 1212  
DB 1201 ATCATCCGCAAC 1212
```

RESULT 8

```
LOCUS E11155 1215 bp DNA linear PAT 29-SEP-1997  
DEFINITION gDNA encoding creatine amidinohydrolase.  
ACCESSION E11155  
VERSION E11155.1 GI:22024796  
KEYWORDS JP 1996089255-A/1.  
SOURCE unclassified  
ORGANISM unclassified.  
REFERENCE 1 (bases 1 to 1215)  
AUTHORS Furukawa, K., Ichikawa, T., Suzuki, M. and Koyama, T.  
TITLE NOVEL, CREATIVE AMIDINOHYDROLASE GENE, NOVEL RECOMBINANT DNA AND  
JOURNAL PRODUCTION OF CREATIVE AMIDINOHYDROLASE  
KIKKOMAN CORP  
Patent: JP 1996089255-A 1 09-APR-1996;  
COMMENT OS Alcaligenes sp. KS-85  
PN JP 1996089255-A/1  
PD 09-APR-1996  
PF 29-SEP-1994 JP 1994235737  
PI FURUKAWA KEISUKE, ICHIKAWA TOSHIO, SUZUKI MASARU, KOYAMA TAIJI  
PC C12N15/09, C12N9/78, (C12N9/78, C12R1:05), (C12N9/78, C12R1:19); CC  
strandedness: Double;  
CC topology: linear;  
FH Key Location/Qualifiers  
FH source 1..1215 /organism="Alcaligenes sp. KS-85" FT CDS  
FT 1..1215 /product="Creatine amidinohydrolase".
```

FEATURES

```
source location/Qualifiers  
1..1215 /organism="unidentified"  
/mol_type="genomic DNA"  
/db_xref="taxon:32644"
```

ORIGIN

```
Query Match      98.0%; Score 1188; DB 6; Length 1215;  
Best Local Similarity 98.8%; Pred. No. 3.1e-136;
```

Matches 1197; Conservative 0; Mismatches 15; Indels 0; Gaps 0;			
Qy	1	ATGACTGACGACATTTTGCAGTGTATGAAATGGCAACACGGGAGAAAGATTATTCGCGG	60
Db	1	ATGACTGACGACATTTTGCAGTGTATGAAATGGCAACACGGGAGAAAGATTATTCGCGG	60
Qy	61	TTTTGGATGCGGAGATACCCGCGCCAAAACAGCTTCGCGGCTGTGATGGCCAAAGAC	120
Db	61	TTTTGGATGCGGAGATACCCGCGCCAAAACAGCTTCGCGGCTGTGATGGCCAAAGAC	120
Qy	121	AATGTCATGCGGCGCTGTCACTTATATCATCTGATCACTACTATATTCGGCTGCTG	180
Db	121	AATGTCATGCGGCGCTGTGTCACTTATATCATCTGATCACTACTATATTCGGCTGCTG	180
Qy	181	TACTGCTATTTTGGACGCAAGTACGGCATGTGATTCAGACCAACAAAGCCACGACGAT	240
Db	181	TACTGCTATTTTGGACGCAAGTACGGCATGTGATTCAGACCAACAAAGCCACGACGAT	240
Qy	241	TGCGCGGCGATCGACGGGGGCGAGCCCTGGCGCGGAGCTTCGGGACAAATCAGCTAC	300
Db	241	TGCGCGGCGATCGACGGGGGCGAGCCCTGGCGCGGAGCTTCGGGACAAATCAGCTAC	300
Qy	301	ACCGACTGGCGCGCGACAAATTTCTATGCGCGCGCTGCGCAGCTGACCAACGCGGCGCAAG	360
Db	301	ACCGACTGGCGCGCGACAAATTTCTATGCGCGCGCTGCGCAGCTGACCAACGCGGCGCAAG	360
Qy	361	CGCATCGGCATCGAGTTGACCAAGTCAATCTGCACTTCGCGCGGCGAGCTGAGAGAAC	420
Db	361	CGCATCGGCATCGAGTTGACCAAGTCAATCTGCACTTCGCGCGGCGAGCTGAGAGAAC	420
Qy	421	CTACCGGGCGTGCAGTTGCTGACATCGACCGCCCTGATGTGATGATGCGGCGCGCG	480
Db	421	CTACCGGGCGTGCAGTTGCTGACATCGACCGCCCTGATGTGATGATGCGGCGCGCG	480
Qy	481	TGCGTGCAGAGACAGAGCTGATTCGCGAGCGCGCGCGCTGTGATGATGCGGCGCGCG	540
Db	481	TGCGTGCAGAGACAGAGCTGATTCGCGAGCGCGCGCGCTGTGATGATGCGGCGCGCG	540
Qy	541	GCCTGCGGCGCTGCGCATCAAGGCGCGCGCTGCGCGCATGATGATGCGGCGCGCG	600
Db	541	GCCTGCGGCGCTGCGCATCAAGGCGCGCGCTGCGCGCATGATGATGCGGCGCGCG	600
Qy	601	AATGCGATGATCGCGGAGATCGCAAAATGCTTCCCTGTGAGAGCTGATGAGACCTGG	660
Db	601	AATGCGATGATCGCGGAGATCGCAAAATGCTTCCCTGTGAGAGCTGATGAGACCTGG	660
Qy	661	ACCTGCTTCACTCGGCGATCAACACCGACGGCGCGCAATCGGCTCAACACCGCATC	720
Db	661	ACCTGCTTCACTCGGCGATCAACACCGACGGCGCGCAATCGGCTCAACACCGCATC	720
Qy	721	GTGCAATCGGCGGACATCTTTGCTGCAACACCTTCCGATGATCTTCGGCTATCTACAC	780
Db	721	GTGCAATCGGCGGACATCTTTGCTGCAACACCTTCCGATGATCTTCGGCTATCTACAC	780
Qy	781	GGGCGTGGAGCGGACGCTGTTCTGCGACATGTCATGAGCGGCGCTGACATCTGGGAG	840
Db	781	GGGCGTGGAGCGGACGCTGTTCTGCGACATGTCATGAGCGGCGCTGACATCTGGGAG	840
Qy	841	AAGAACTGGCCGATCGCGCGGCGCTGAGCTGATCAAGCCGGGCGCGGCTGGCAG	900
Db	841	AAGAACTGGCCGATCGCGCGGCGCTGAGCTGATCAAGCCGGGCGCGGCTGGCAG	900
Qy	901	GACATCGGCATCGAGTCAACAGATGTACCGGAGTGGGACCTGTGAAAGCCGCTC	960
Db	901	GACATCGGCATCGAGTCAACAGATGTACCGGAGTGGGACCTGTGAAAGCCGCTC	960
Qy	961	TTGCGCTAGTGGCACTCTTCGCGCGTGTGCTGCACTTACGCTGCGAGCGGCGG	1020
Db	961	TTGCGCTAGTGGCACTCTTCGCGCGTGTGCTGCACTTACGCTGCGAGCGGCGG	1020
Qy	1021	GAGCTGCGGAGACATCGACACCGAGCTGAAGCCGGGAGTGGTCTTCATGAGCGG	1080
Db	1021	GAGCTGCGGAGACATCGACACCGAGCTGAAGCCGGGAGTGGTCTTCATGAGCGG	1080

Qy	1081	ATGCTGATGCTGCCGAGAGGCGATCCCGGTGCTTCGCGGCTATTCGAGACAGACATCTCG	1140
Db	1081	ATGCTGATGCTGCCGAGAGGCGATCCCGGTGCTTCGCGGCTATTCGAGACAGACATCTCG	1140
Qy	1141	ATCGTGGGAGGACCGTTCGCGAACAATCACCGGCTTCCGTTTCGCTCCGGAACAGAC	1200
Db	1141	ATCGTGGGAGGACCGTTCGCGAACAATCACCGGCTTCCGTTTCGCTCCGGAACAGAC	1200
Qy	1201	ATCATCGGCAAC 1212	
Db	1201	ATCATCGGCAAC 1212	
RESULT 9			
LOCUS	AX721933	1215 bp	DNA linear PAT 07-MAY-2003
DEFINITION	Sequence 1 from Patent EP1298213.		
ACCESSION	AX721933		
VERSION	AX721933.1	GI:30422515	
KEYWORDS			
SOURCE	Erwinia sp.		
ORGANISM	Erwinia sp.		
REFERENCE	Shao, Z., Schmuck, R., Kratzsch, P., Kenkiles, J. and Weisser, H.		
AUTHORS	Erwinia sp.		
TITLE	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;		
JOURNAL	Enterobacteriaceae; Erwinia.		
FEATURES	1		
source	location/Qualifiers		
	1..1215		
	/organism="Erwinia sp."		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:558"		
	1..1215		
	/note="unassigned protein product"		
	/codon_start=1		
	/transl_table=1		
	/db_xref="GI:30422515"		
	/translation="MTDDMLHVMKNGHNGEKYSPSPDAEMTRROSDVPRMAENDVDA		
	ALFTSYHCINYSGLFYCFGRKGYVIDDHAATITISAGITGGQMRRSFDNITTYTD		
	WRDNFYQAVRLTPGARIGIEPDHVLDPRLTLEALPVEFVDIGQPMRMSFDNITTYTD		
	SLEROKIIRSGARICDVGAACVAAGVHEHVAIATTNAAVEIRLXSPFVELMD		
	TWTFQSGINDGHNPTNRIVSGDILSNTPPMI FGYYTALER TLFCHEVDASL		
	DYERKNVAVRHRIELIKPGARCDIATLEMTREMDILKYSRSGVSHSY		
	GREAGVELREDITVLDPGMVSMPEPMVLPFGAPGAGYREHDLITVGEDEAENITG		
	PFPGREHNIIRN"		
ORIGIN			
Query Match	81.3%; Score 984.8; DB 6; Length 1215;		
Best Local Similarity	88.3%; Pred. No. 2.1e-111;		
Matches 1070; Conservative 0; Mismatches 142; Indels 0; Gaps 0;			
Qy	1	ATGACTGACGACATTTTGCAGTGTATGAAATGGCAACACGGGAGAAAGATTATTCGCGG	60
Db	1	ATGACTGACGACATTTTGCAGTGTATGAAATGGCAACACGGGAGAAAGATTATTCGCGG	60
Qy	61	TTTTGGATGCGGAGATACCCGCGCCAAAACAGCTTCGCGGCTGTGATGGCCAAAGAC	120
Db	61	TTTTGGATGCGGAGATACCCGCGCCAAAACAGCTTCGCGGCTGTGATGGCCAAAGAC	120
Qy	121	AATGTCATGCGGCGCTGTCACTTATATCATCTGATCACTACTATATTCGGCTGCTG	180
Db	121	AATGTCATGCGGCGCTGTGTCACTTATATCATCTGATCACTACTATATTCGGCTGCTG	180
Qy	181	TACTGCTATTTTGGACGCAAGTACGGCATGTGATTCAGACCAACAAAGCCACGACGAT	240
Db	181	TACTGCTATTTTGGACGCAAGTACGGCATGTGATTCAGACCAACAAAGCCACGACGAT	240
Qy	241	TGCGCGGCGATCGACGGGGGCGAGCCCTGGCGCGGAGCTTCGGGACAAATCAGCTAC	300
Db	241	TGCGCGGCGATCGACGGGGGCGAGCCCTGGCGCGGAGCTTCGGGACAAATCAGCTAC	300

```

Db 241 TCGCCGCGCATGATGAGCGGTGACGCCCTGCGCGCCGTGAGCTTCCGCGAACAATCACTAT 300
Qy 301 ACCGATCGGCGCGCGGACAAATTTTATCGCGCGCGTGGCGAGCTGACCAACGCGGCGCCAG 360
Db 301 ACCGATCGGCGCGCGGACAACTTTTACCAAGCGCGCGCTCCCACTCAACCCCGCGCGCCAG 360
Qy 361 CGCATCGGCGCATGAGTTGACCAACGTCATCTGCACTTCCGCGCGCAGCTCGAGGAGCC 420
Db 361 CGCATCGGCGCATGAGTTGACCAACGTCATCTGCACTTCCGCGCGCAGCTCGAGGAGCG 420
Qy 421 CTACCGGCGCGTGAAGTTGTCGACATCAAGCCAGCCCTGATGATGAGTGGCACTCAAG 480
Db 421 CTGCGCGCGCGTGAAGTTGTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 480
Qy 481 TCGCTCGAAGAGCAAGCTGATTCGCGAAGCGCGCGCGTGTGATGAGTGGCGCGCGCG 540
Db 481 TCGCTCGAAGAGCAAGCTGATTCGCGAAGCGCGCGCGTGTGATGAGTGGCGCGCGCG 540
Qy 541 GCTGCGCGCGTGCATCAAGCGCGCGTGCAGCGAGTGAAGTGGCGATCGCCACCAAC 600
Db 541 GCTGCGCGTTCGCGCGCGTGCAGCGCGCGTTCGCGAGCAAGAGTGGCGATCGCCACCAAC 600
Qy 601 AATGCGATGATTCGCGAGATTCGCAATGCTTCCCTTCTGAGAGCTGATGAGCACTGG 660
Db 601 AATGCGATGATTCGCGAGATTCGCAATGCTTCCCTTCTGAGAGCTGATGAGCACTGG 660
Qy 661 ACCTGCTTCAGTGGGCGATCAACAGCGCGCGCGCAATCGGATCAACCAACCGCATC 720
Db 661 ACCTGCTTCAGTGGGCGATCAACAGCGCGCGCGCAATCGGATCAACCAACCGCATC 720
Qy 721 GTGCAATTCGCGCGCATCTCTTCTGCAACAACCTTCCGATGATCTTCCGCTACTACAC 780
Db 721 GTGCAATTCGCGCGCATCTCTTCTGCAACAACCTTCCGATGATCTTCCGCTACTACAC 780
Qy 781 GCGCTGAGCGCAAGCTGTTCTGCGACCAATGTCATGACGCCAGCTTGAATCTGGAG 840
Db 781 GCGCTGAGCGCAAGCTGTTCTGCGACCAATGTCATGACGCCAGCTTGAATCTGGAG 840
Qy 841 AAGAACGTGGCGGTGATCGCGCGCGCGTTCGAGCTGATCAAGCGCGCGCGCTGCAAG 900
Db 841 AAGAACGTGGCGGTGATCGCGCGCGCGTTCGAGCTGATCAAGCGCGCGCGCTGCAAG 900
Qy 901 GACATCGCGCATGAGCTCAACAGAGTATACCGCGAGTGGAGACTTGTCAAGTACCGCTCC 960
Db 901 GATATCGCGCATGAGCTCAACAGAGTATACCGCGAGTGGAGACTTGTCAAGTACCGCTCC 960
Qy 961 TTGCGCTATGAGCACTCTTCTGCGCGTGTGTCATCACTACGATCGCGCGCGCGCTG 1020
Db 961 TTGCGCTATGAGCACTCTTCTGCGCGTGTGTCATCACTACGATCGCGCGCGCGCTG 1020
Qy 1021 GAGCTGCGCGAGGACATGACACCGAGTGAAGCCCGGCGATGATGATGATGATGATGATG 1080
Db 1021 GAGCTGCGCGAGGACATGACACCGAGTGAAGCCCGGCGATGATGATGATGATGATGATG 1080
Qy 1081 ATGCGATGCTGCGCGAGGAGTCCCGGAGTGGCGCGGCTATCGAGGACACGATCTCTG 1140
Db 1081 ATGCGATGCTGCGCGAGGAGTCCCGGAGTGGCGCGGCTATCGAGGACACGATCTCTG 1140
Qy 1141 ATGCGTGGGAGAGAGCGTGTCCGAGAACATCAACGCGCTTCCGCTTGGTTCGAGAACAC 1200
Db 1141 ATGCGTGGGAGAGAGCGTGTCCGAGAACATCAACGCGCTTCCGCTTGGTTCGAGAACAC 1200
Qy 1201 ATCATCGCGAAC 1212
Db 1201 ATCATCGCGAAC 1212

```

```

RESULT 10
AX721949 1215 bp DNA linear PART 07-MAY-2003
LOCUS AX721949 Sequence 17 from Patent EP1298213.
DEFINITION AX721949
ACCESSION AX721949
VERSION AX721949.1 GI:30422525

```

```

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
    source
        1..1215
            /organism="synthetic construct"
            /mol_type="unassigned DNA"
            /db_xref="taxon:32630"
            /note="variant C1gc2"
            1..1215
                /note="unassigned protein product"
                /codon_start=1
                /protein_id="CAD90221.1"
                /db_xref="gi:30422526"
                /translation="MTDDMLHVMKMHNGEKESPPSDAEMTRROSDVVRMMANEDVDA
                ALPTSHCINITYSGMICYRGRKGMVITDHAATTSAGTGGCPMPRSFGDNTITYTD
                WRDNFYQAVRQILTPGARIGIEFDHYDLDPRFLIEALPEYEPVDICQPSMRETVK
                SLERQKILREGARICDVGAACVAAVAGVHEHVALATNAMAREINKSPFVELMD
                TWTFQSGINTDGMNPVTRIVQSGDILSNTPPMIPIGYTALERTLPCDHVDASL
                DTYEKNTAVHRRGIELIKPGARCDIAIENLMEHREMDLLKRSFGLYSGVLSHY
                GREAGVELREDIDTVLPQMVVSMPEVMPVMEPGAAGAGYREHDLIVGEIDAEVITG
                PPEPBNIRIN"
    CDS
        1..1215
            /note="unassigned protein product"
            /codon_start=1
            /protein_id="CAD90221.1"
            /db_xref="gi:30422526"
            /translation="MTDDMLHVMKMHNGEKESPPSDAEMTRROSDVVRMMANEDVDA
            ALPTSHCINITYSGMICYRGRKGMVITDHAATTSAGTGGCPMPRSFGDNTITYTD
            WRDNFYQAVRQILTPGARIGIEFDHYDLDPRFLIEALPEYEPVDICQPSMRETVK
            SLERQKILREGARICDVGAACVAAVAGVHEHVALATNAMAREINKSPFVELMD
            TWTFQSGINTDGMNPVTRIVQSGDILSNTPPMIPIGYTALERTLPCDHVDASL
            DTYEKNTAVHRRGIELIKPGARCDIAIENLMEHREMDLLKRSFGLYSGVLSHY
            GREAGVELREDIDTVLPQMVVSMPEVMPVMEPGAAGAGYREHDLIVGEIDAEVITG
            PPEPBNIRIN"
    ORIGIN
        Query Match 81.3%; Score 984.8; DB 6; Length 1215;
        Best Local Similarity 88.3%; Pred. No. 2.1e-111;
        Matches 1070; Conservative 0; Mismatches 142; Indels 0; Gaps 0;
        1 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        2 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        3 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        4 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        5 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        6 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        7 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        8 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        9 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        10 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        11 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        12 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        13 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        14 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        15 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        16 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        17 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        18 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        19 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        20 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        21 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        22 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        23 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        24 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        25 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        26 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        27 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        28 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        29 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        30 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        31 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        32 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        33 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        34 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        35 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        36 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        37 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        38 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        39 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        40 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        41 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        42 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        43 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        44 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        45 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        46 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        47 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        48 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        49 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        50 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        51 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        52 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        53 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        54 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        55 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        56 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        57 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        58 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        59 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60
        60 ATGACTGACGACATGTTGCACTGATGAAATGCGCAACGCGGAGAAATTAATTCGCG 60

```

Db 541 GCCCTGCTTCCCGCTGCAAGCCGCGCTTCCGAGACGACGAGTCCGATCCGACACC 600
 QY 601 AATGCGATGATCCGCGAGATGCGCAATGCTTCCCTGCTGAGGCTGATGACACCTGG 660
 Db 601 AATGCGATGATCCGCGAGATGCGCAATGCTTCCCTGCTGAGGCTGATGACACCTGG 660
 QY 661 ACCCTGCTTCCAGTCCGCGCATCAACACGCGCGCGACATCCGCTCAACACCGCATC 720
 Db 661 ACCCTGCTTCCAGTCCGCGCATCAACACGCGCGCGACATCCGCTCAACACCGCATC 720
 QY 721 GTGCAATCCGCGCGCATCTCTTCTGCTCAACACTTCCCGATGATCTTCCGCTCAAC 780
 Db 721 GTGCAATCCGCGCGCATCTCTTCTGCTCAACACTTCCCGATGATCTTCCGCTCAAC 780
 QY 781 GCGCTGAGAGCGACGCTGTTCTGCGACCATGATGATGACGCGCTCGACATCTGAGAG 840
 Db 781 GCGCTGAGAGCGACGCTGTTCTGCGACCATGATGATGACGCGCTCGACATCTGAGAG 840
 QY 841 AAGAACTGCGCGCTGATGCGCGCGCGCTGAGCTGATCAAGCCGCGCGCGCTGAGAG 900
 Db 841 AAGAACTGCGCGCTGATGCGCGCGCGCTGAGCTGATCAAGCCGCGCGCGCTGAGAG 900
 QY 901 GACATGCGCATGCAAGCTCAACGAGATGACCGGAGTGGGATCTGCTGAAGTACCGCTCC 960
 Db 901 GATATGCGCATGCAAGCTCAACGAGATGACCGGAGTGGGATCTGCTGAAGTACCGCTCC 960
 QY 961 TTGCGCTATGCGCGCTGCTTCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
 Db 961 TTGCGCTATGCGCGCTGCTTCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
 QY 1021 GAGCTGCGCGAGGACATCAACGAGCTGAGACCGCGCGCTGATGCTTCAATGAGCGG 1080
 Db 1021 GAGCTGCGCGAGGACATCAACGAGCTGAGACCGCGCGCTGATGCTTCAATGAGCGG 1080
 QY 1081 ATGCTGATCTGCTGCGAGGCGATCCCGCTGCGCGCTGATGCTGCGAGCACTGCTG 1140
 Db 1081 ATGCTGATCTGCTGCGAGGCGATCCCGCTGCGCGCTGATGCTGCGAGCACTGCTG 1140
 QY 1141 ATGCTGCGGAGGAGCGCTGCGGAGACATCAACGCGCTTCCCGCTGCGGAGCAAC 1200
 Db 1141 ATGCTGCGGAGGAGCGCTGCGGAGACATCAACGCGCTTCCCGCTGCGGAGCAAC 1200
 QY 1201 ATCATCCGCAAC 1212
 Db 1201 ATCATCCGCAAC 1212

RESULT 11
 AX721941 1215 bp DNA linear PART 07-MAY-2003
 LOCUS Sequence 9 from Patent BP1298213.
 DEFINITION AX721941
 ACCESSION AX721941.1 GI:30422517
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 synthetic construct
 synthetic construct
 other sequences; artificial sequences.
 REFERENCE
 1 Shao, Z., Schumuck, R., Kratzsch, P., Kenklies, J. and Weisner, H.
 AUTHORS Variants of an erwinia-type creatinase
 TITLE Patent: EP 1298213-A 9 02-APR-2003;
 JOURNAL Roche Diagnostics GmbH (DE) / F.HOFMANN-LA ROCHE AG (CH)
 FEATURES
 location/Qualifiers
 1..1215
 /organism="synthetic construct"
 /mol_type="unassigned DNA"
 /db_xref="taxon:32630"
 /note="variant CT1m24"
 1..1215
 /note="unnamed protein product"
 /codon_start=1
 /transl_table=11
 /protein_id="CAD90217.1"

ORIGIN

Query Match 81.0%; Score 981.6; DB 6; Length 1215;
 Best Local Similarity 88.1%; Pred. No. 5.2e-111;
 Matches 1068; Conservative 0; Mismatches 144; Indels 0; Gaps 0;

/db_xref="GI:30422518"
 /translation="MTDMLHVMKHNKSEKYSPESDAEMTRROSQVRRMAENDVDA
 ALFTSYHCINYSGLFYCGFKYGMVINDODHATITSGIDGQPMRSGFNITFTD
 WRDNPYQAVNLTGPGARIGIEPHVDLTPERTLEBALRGYEPVDIGQSPMMRTVK
 SLEBOKLIREGRKICDVGACVAAVKAGVPEHEVALATTNMAWBEIASFPEVLEMD
 TWTFQSGINTDGANPVTNRIVQSGDILSLTFTPEYVTALEHTLDCDHVDSL
 DTWEKVAIVHRGDELIRGACKDIAELEMYEBMDLTKTRSRFGYGSFGLSHY
 GSEAGVEIREDDIDVTIQQPMVSMPEMPLPEGAPGAGGVRHDIILIVGDEAENITG
 PFGPEHNIIRN"

QY 1 AATGCTGACGACATGCGACGTGATGAAATGGCAACCGCGAGAAAGATTTCGCGG 60
 Db 1 AATGCTGACGACATGCGACGTGATGAAATGGCAACCGCGAGAAAGATTTCGCGG 60
 QY 61 TTTTGGATGCGGAGATGACCGCGCGCAAAACGATTCGCGCTGATGCGCAAGAC 120
 Db 61 TTTTGGATGCGGAGATGACCGCGCGCAAGATGACGCGCGCTGATGCGCAAGAC 120
 QY 121 AATGCTGATGCGGCGCTGTTCACTCTTATCACTGCACTCACTATTTCGCGCTG 180
 Db 121 GACGTGACGCTGCGCTGTTCACTCTTATCACTGCACTCACTATTTCGCGCTG 180
 QY 181 TACTGCTATTTTGGACGCAAGTACGCGCATGCTGATGACCAACACCGCGACGAT 240
 Db 181 TACTGCTATTTTGGCGGCAAAATACGCGCATGCTGATGACCAAGACCATGCGAC 240
 QY 241 TCGCGCGGATGACG 300
 Db 241 TCGCGCGGATGACG 300
 QY 301 ACCGACTGCGCGCGCGCAATTTCTATGCGCGCTGCGCGCACTGACCAAGGCGCG 360
 Db 301 ACCGACTGCGCGCGCGCAATTTCTATGCGCGCTGCGCGCACTGACCAAGGCGCG 360
 QY 361 CGCATGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 420
 Db 361 CGCATGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 420
 QY 421 CTACCGGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
 Db 421 CTACCGGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
 QY 481 TCGCTGAAAGACAGACCTGATCCGCAAGGCGCGCGCTGATGATGATGATGATG 540
 Db 481 TCGCTGAAAGACAGACCTGATCCGCAAGGCGCGCGCTGATGATGATGATGATG 540
 QY 541 GCGTTCGCGCGCGCGCATGAAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
 Db 541 GCGTTCGCGCGCGCGCATGAAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
 QY 601 AATGCGATGATCCGCGAGATGCGCAATGCTTCCCTTCTGCTGAGAGCTGAGAC 660
 Db 601 AATGCGATGATCCGCGAGATGCGCAATGCTTCCCTTCTGCTGAGAGCTGAGAC 660
 QY 661 ACCCTGCTTCCAGTCCGCGCATCAACACGCGCGCGCGCAATCCGCTGACCAAC 720
 Db 661 ACCCTGCTTCCAGTCCGCGCATCAACACGCGCGCGCGCAATCCGCTGACCAAC 720
 QY 721 GTGCAATCCGCGCGCATCTCTTCTGCTCAACACTTCCCGATGATCTTCCGCTCA 780
 Db 721 GTGCAATCCGCGCGCATCTCTTCTGCTCAACACTTCCCGATGATCTTCCGCTCA 780
 QY 781 GCGCTGAGAGCGACGCTGTTCTGCGACCATGATGATGACGCGCTCGACATCTGAG 840
 Db 781 GCGCTGAGAGCGACGCTGTTCTGCGACCATGATGATGACGCGCTCGACATCTGAG 840
 QY 841 AAGAACTGCGCGCTGATGCGCGCGCGCTGAGCTGATCAAGCCGCGCGCGCTGAG 900

Db 841 AAGACGTCGCGCGTGCACCGCGCGGCTCGAACTCATCAAGCCGGGTGCGCGCTGCAG 900
Qy 901 GACATGCGCATTCAGCTGCACAGAGATGTACCGGAGGAGGAGCTGCTGAATACCGCTCC 960
Db 901 GATATGCGCATTCAGCTGCACAGAGATGTACCGGAGGAGGAGCTGCTGAATACCGCTCC 960
Qy 961 TTCGGCTATGCGCATTCCTTCCGCGTGTGCTGCACTACTACGATCGCGAGCGCGCTG 1020
Db 961 TTCGGCTATGCGCATTCCTTCCGCGTGTGCTGCACTACTACGATCGCGAGCGCGCTG 1020
Qy 1021 GAGCTGCGGAGGAGCATTCGACACCGAGCTGAAAGCCGCGATGCTGTCTCATGAGCCG 1080
Db 1021 GAACTGCGGAGGAGCATTCGATCCGCTGCTGACGCGCGATGCTGTCTCATGAGCCG 1080
Qy 1081 ATGCTATGCTGCGGAGGAGCATTCGCGGTGCGCGCGCTATGCGGAGGAGCATTCCTG 1140
Db 1081 ATGCTATGCTGCGGAGGAGCATTCGCGGTGCGCGCGCTATGCGGAGGAGCATTCCTG 1140
Qy 1141 ATGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200
Db 1141 ATGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200
Qy 1201 ATCATCGGCAAC 1212
Db 1201 ATCATCGGCAAC 1212

RESULT 12
AX721943 1212 bp DNA linear PAT 07-MAY-2003
LOCUS AX721943
DEFINITION Sequence 11 from Patent EP1298213.
ACCESSION AX721943
VERSION AX721943.1 GI:30422519
KEYWORDS
SOURCE
ORGANISM
other sequences; artificial sequences.
REFERENCE
1 Shao, Z., Schmuck, R., Kratzsch, P., Kenkies, J. and Weisner, H.
Variants of an erwinia-type creatinase
Patent: EP 1298213-A 11 02-APR-2003;
Roche Diagnostics GmbH (DE) ; F.HOFFMANN-LA ROCHE AG (CH)
LOCATION/Qualifiers
FEATURES
source
1.1212
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="variant CT2m9"
1..>1212
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAD90218.1"
/db_xref="GI:30422520"
/translation="MTDDMLHVMKMHNGEKEYSPFSDAENTRQSDVRRMAMENDVDA
ALPTSHCHINYSGRPLCYGRKGMVINDHATTISAGIDGQSPRRSRGNDITTD
WRDNFTQAVRQLTPGARRTGIEFDHVDLFRITLSEALPGFVADIGOSPMMRTVK
SLEBQQLIREGARIQVNGACVAAVKAGPEHEVALATTNAVERILASFPVELSLD
TWTWPGSINTIDANHPVTRIVOSGDIISLNTFPMIFGYTALBERTLFCDDHDDSL
DTWEKNAVHRRGLELTKPGARKDIALIENEMYREMDLTKYSPGSHFVLSHY
GREGAVELREDIDITVALQPMVVSMEPMVMPEGAPGAGYREHDIIVGDDAENITG
PFPEPHNITRN"

ORIGIN
Query Match 80.9%; Score 980; DB 6; Length 1212;
Best Local Similarity 88.0%; Pred. No. 8.2e-111;
Matches 1067; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

Qy 1 ATGACTGAGCAGCATGTTGCACTGATGAATGGCAACGCGGAGAAAGTTATTGCGCG 60
Db 1 ATGACTGAGCAGCATGTTGCACTGATGAATGGCAACGCGGAGAAATATTCGCCCG 60
Qy 61 TTTTGGATGCCGAGATGACCGCGCGCAAAAGAGCTTGGCGGCTGAGATGGCAAGAAC 120

Db 61 TTTTCCGATGCCGAGATGACGCGCGCGCAGAGTGAAGTGGCGGCTGAGATGGCGCAAAAC 120
Qy 121 AATGTGAGATGCCGCGCTGTTTCACTCTTATCATGTGATCAATCACTATTTCGGCTGCTG 180
Db 121 GAGGTGAGCGCTGCGTGTTCACCTCTCATATTGATGATCACTACTACTGATTCCTG 180
Qy 181 TACTGCTATTTTGGAGCGCAAGTACGAGCATGATGATGATGATGATGATGATGATGATGAT 240
Db 181 TACTGCTATTTTGGAGCGCAAAATACGAGCATGATGATGATGATGATGATGATGATGATGAT 240
Qy 241 TCGGCGGAGATGACGCGCGCGAGCGCTGCGCGCGCGAGCTTGGCGCAAAATCACTAC 300
Db 241 TCGGCGGAGATGATGACGCGCGCGAGCGCTGCGCGCGCGAGCTTGGCGCAAAATCACTAC 300
Qy 301 ACCGATGCGCGCGCGAGCAATTTTCTATTCGCGCGCGCGCGAGCTTGGCGCAAAATCACT 360
Db 301 ACCGATGCGCGCGCGCGAGCAATTTTCTATTCGCGCGCGCGCGAGCTTGGCGCAAAATCACT 360
Qy 361 CGCATGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
Db 361 CGCATGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
Qy 421 CTACCGGCGGCTGAGTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
Db 421 CTACCGGCGGCTGAGTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
Qy 481 TCGCTGAG 540
Db 481 TCGCTGAG 540
Qy 541 GCGTGGCGGCTGCGCATGACAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
Db 541 GCGTGGCGGCTGCGCATGACAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
Qy 601 AATGCGATGATTCGCGAGATTCGCGAAATGTTCCCTTCTGCTGAGCTGATGACACTG 660
Db 601 AATGCGATGATTCGCGAGATTCGCGAAATGTTCCCTTCTGCTGAGCTGATGACACTG 660
Qy 661 ACTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
Db 661 ACTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
Qy 721 GTGCAATCGGCGAGATTCCTTCTGCTGCAACCTTCCGATGATCTTCCGCTACTACAC 780
Db 721 GTGCAATCGGCGAGATTCCTTCTGCTGCAACCTTCCGATGATCTTCCGCTACTACAC 780
Qy 781 GCGCTGAG 840
Db 781 GCGCTGAG 840
Qy 841 AAGACGTCGCGCGTGCATTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
Db 841 AAGACGTCGCGCGTGCATTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
Qy 901 GACATGCGCATGAGCTGCAACGAGATGTAACCGGAGTGGGAGCTGCTGAAGTACCGCTCC 960
Db 901 GACATGCGCATGAGCTGCAACGAGATGTAACCGGAGTGGGAGCTGCTGAAGTACCGCTCC 960
Qy 961 TTCGGCTATGCGCATTCCTTCCGCGTGTGCTGCACTACTACGATCGCGAGCGCGCTG 1020
Db 961 TTCGGCTATGCGCATTCCTTCCGCGTGTGCTGCACTACTACGATCGCGAGCGCGCTG 1020
Qy 1021 GAGCTGCGGAGGAGCATTCGACACCGAGCTGAAAGCCGCGATGCTGTCTCATGAGCCG 1080
Db 1021 GAACTGCGGAGGAGCATTCGATCCGCTGCTGACGCGCGATGCTGTCTCATGAGCCG 1080
Qy 1081 ATGCTATGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
Db 1081 ATGCTATGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
Qy 1141 ATGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200

Db 1141 ATGCTGGGGAGACCGCGCAGAGAAATTAACCGATTCCCTTCCGGCGCTGAGACAAAC 1200
QY 1201 ATCATCCGCAAC 1212
Db 1201 ATCATCCGCAAC 1212
RESULT 13
AX721945 1212 bp DNA linear PAT 07-MAY-2003
DEFINITION Sequence 13 from Patent EP1298213.
ACCESSION AX721945
VERSION AX721945.1 GI:30422521
KEYWORDS
SOURCE
ORGANISM
synthetic construct
synthetic construct
other sequences; artificial sequences.
REFERENCE
AUTHORS Shao, Z., Schmuck, R., Kratzsch, P., Kenlles, J. and Weisner, H.
TITLE Variants of an erwinia-type creatinase
JOURNAL Patent: EP 1298213-A 13 02-APR-2003;
Roche Diagnostics GmbH (DE) ; F.HOFFMANN-LA ROCHE AG (CH)
FEATURES
source
1. 1212
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="variant CT2m10"
1. >1212
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAD90219.1"
/db_xref="GI:30422522"
/translation="MTDMLHYMKHNGKEYSPSDAMETROSQVRRMAENDVDA
ALPTSHCHINYSGLPYCYGFRKGVVDDODHATTSAGIDGGQPMRSPGNITYTD
WRDNPYQVROLTPGARIRIGEPHVDLPRTTLEBALPGEPIDGSPMMWETVK
SLEBCLIREGARICDVGACAAVAAKAVPHEVAIATINMAVEIASFPFVELMD
TWTWQSGINTGANTLPVNRIVQSGDILSLNTPFMI FGYYALERTYFCDDVDASL
DTWEKVAVHRRGLNLIKPGARCKDIALINEMTREMDDLKTYRSFGYHSFVLSHY
GREAGVELREDIDTVLQPGMVVSMEMPVWLPGAGPAGGYREHDLIVGEDAENITG
PLPGPHNHTIRN"
ORIGIN
Query Match 80.9%; Score 980; DB 6; Length 1212;
Best Local Similarity 88.0%; Pred. No. 8.2e-111;
Matches 1067; Conservative 0; Mismatches 145; Indels 0; Gaps 0;
QY 1 ANGACTGAGCAATGTTTGCACGTGATGAATAAGGCAACAGGCGAAGAAATTTATTCGCG 60
Db 1 ATGACTGAGCAATGTTTGCACGTGATGAATAAGGCAACAGGCGAAGAAATTTATTCGCG 60
QY 61 TTTTGGAGATCGGAGATGACCGCGCGCAAAACGAGCTTCGCGGCTGATGCGCAAGAC 120
Db 61 TTTTCCGATGCGAGATGACCGCGCGCAAGAGAGCTTCGCGGCTGATGCGCAAGAC 120
QY 121 AATGTCGATGCGCGCTGTTCACTCTTATCACTGATCAACTACTATTCCGGCTGGCTG 180
Db 121 GAGCTGAGCGCTGCGCTGTTCACTCTTATCACTGATCAACTACTATTCCGGCTGGCTG 180
QY 181 TACTGCTATTTTGGAGCGCAAGTACGCGATGTCATATGACCAACAAAGCCGACGAGATT 240
Db 181 TACTGCTATTTTGGAGCGCAAGTACGCGATGTCATATGACCAACAAAGCCGACGAGATT 240
QY 241 TCGGCGCGGATGACCGCGCGCGAGCTTCGCGGCTGCGGCAACATCACTAC 300
Db 241 TCGGCGCGGATGACCGCGCGCGAGCTTCGCGGCTGCGGCAACATCACTAC 300
QY 301 ACCGACTGGCGCGCGCAAAATTTCTATCGCGCGCTGCGCGCAAGTACCAAGCGCGCAAG 360
Db 301 ACCGACTGGCGCGCGCAAAATTTCTATCGCGCGCTGCGCGCAAGTACCAAGCGCGCAAG 360
QY 361 CGCATCGCATGAGTTTCCAGCAACGTCATCTCGACTTCGCGCGCGCAAGTCCAGAGAGCC 420

Db 361 CGCATCGCATGAGTTTCCAGCAACGTCATCTCGACTTCGCGCGCGCAAGTCCAGAGAGCC 420
QY 421 CTACCGGCGCTGAGATTGTTGACATCAAGCCAGCCCTGATGTGATGCGACATCAAG 480
Db 421 CTGCGCGCGCTGAGATTGTTGACATCAAGCCAGCCCTGATGTGATGCGACATCAAG 480
QY 481 TCGCTGAAAGAGCAAGACTGATCCGGAAGCGCGCGCTGATGTGACATGCGCGCGCG 540
Db 481 TCGCTGAAAGAGCAAGACTGATCCGGAAGCGCGCGCTGATGTGACATGCGCGCGCG 540
QY 541 GCTTCGCGGCTGCTCATCAAGCGCGCGCTGCGCGCAAGTGAAGTGGAGATGCGCAACC 600
Db 541 GCTTCGCGGCTGCTCATCAAGCGCGCGCTGCGCGCAAGTGAAGTGGAGATGCGCAACC 600
QY 601 AATGCGATGATCCGCGAGATGCGCAATTCCTTCCCTTCTGAGACTGATGACACTG 660
Db 601 AATGCGATGATCCGCGAGATGCGCAATTCCTTCCCTTCTGAGACTGATGACACTG 660
QY 661 ACCTGGTTCAGTGGGATCAACACGAGCGCGCGCAATTCGCGTCAACACCGCATC 720
Db 661 ACCTGGTTCAGTGGGATCAACACGAGCGCGCGCAATTCGCGTCAACACCGCATC 720
QY 721 GTGCAATCCGCGGACATCTTTCGCTCAACACTTCCGATGATCTTCCGCTACTACAC 780
Db 721 GTGCAATCCGCGGACATCTTTCGCTCAACACTTCCGATGATCTTCCGCTACTACAC 780
QY 781 GCGCTGAGCGCAAGCTGTTCTGCGACATGTCATGACCGCAAGCTGACATCTGGAG 840
Db 781 GCGCTGAGCGCAAGCTGTTCTGCGACATGTCATGACCGCAAGCTGACATCTGGAG 840
QY 841 AAGAACGTGGCGCTGCTGATGCGCGCGGCTGCGAGCTGATCAACGCGCGCGCGCTGCAAG 900
Db 841 AAGAACGTGGCGCTGCTGATGCGCGCGGCTGCGAGCTGATCAACGCGCGCGCGCTGCAAG 900
QY 901 GACATCGCATGAGCTCAACAGAGTGTACCGCGAGTGGACCTGCTGAAGTACCGCTCC 960
Db 901 GATATCGCATGAGCTCAACAGAGTGTACCGCGAGTGGACCTGCTGAAGTACCGCTCC 960
QY 961 TTGGCTATGCGCACTCTTTCGCGCTGCTGTCACATCACTACGATGTCGAGCGCGCGCTG 1020
Db 961 TTGGCTATGCGCACTCTTTCGCGCTGCTGTCACATCACTACGATGTCGAGCGCGCGCTG 1020
QY 1021 GAGCTGCGCGAGATGATGACACCGAGCTGAAGCGCGCGCATGAGTCTCAATGAGCGG 1080
Db 1021 GAGCTGCGCGAGATGATGACACCGAGCTGAAGCGCGCGCATGAGTCTCAATGAGCGG 1080
QY 1081 ATGATGATGCTGCGCGAGCGCATGCGCGGCTGCGCGCTATGCGAGACGACATCTCG 1140
Db 1081 ATGATGATGCTGCGCGAGCGCATGCGCGGCTGCGCGCTATGCGAGACGACATCTCG 1140
QY 1141 ATGCTGGGAGAGAGCGGTGCGAGAAATCAACCGGCTTCCGTTGGTTCGGAACACAAAC 1200
Db 1141 ATGCTGGGAGAGAGCGGTGCGAGAAATCAACCGGCTTCCGTTGGTTCGGAACACAAAC 1200
QY 1201 ATCATCCGCAAC 1212
Db 1201 ATCATCCGCAAC 1212
RESULT 14
AX721947 1215 bp DNA linear PAT 07-MAY-2003
DEFINITION Sequence 15 from Patent EP1298213.
ACCESSION AX721947
VERSION AX721947.1 GI:30422523
KEYWORDS
SOURCE
ORGANISM
synthetic construct
synthetic construct
other sequences; artificial sequences.
REFERENCE
AUTHORS Shao, Z., Schmuck, R., Kratzsch, P., Kenlles, J. and Weisner, H.
TITLE Variants of an erwinia-type creatinase

JOURNAL Patent: EP 1298213-A 15 02-APR-2003;
Roche Diagnostics GmbH (DE) ; F.HOFFMANN-LA ROCHE AG (CH)
FEATURES Location/Qualifiers

source
1..1215
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="variant CT2m28"
1..1215
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAD90220.1"
/db_xref="GI:30422524"
/translation="MTDMLHVMKHNKEKYSPESDAEMTRPSDVRMAENDVDA
ALFTSYHCINYSGLICYTFRKRYGAVITDQHAATTISAGIDGQPRRSFGNITTTD
WRNDYQAVRQLTPGARRIIGLEFHDVLDPEFRTLEBALPGVEFVIGPSPMWRIVK
SLEBQCLIREGARI CDVGGA CVAVKAGVPEHEVAIATNVAVEIASKSPFVELMD
TWTWFSQINTDGAHPVTRNI VOSGDILSLNTPMI PGYTALERTLCFDHVDASL
DTWEKNVAVHRRLGELIKPGARKDIALELNEMRYEMDLKXRSFGYGSFGVLSHY
GRBAGVELREDIDTVLQPGMVSMEMWMLPEGA PGAGYREHDLIIVGDEAGNITG
PFGPEPHNITRN"

ORIGIN

Query Match 80.9%; Score 980; DB 6; Length 1215;
Best Local Similarity 88.0%; Pred. No. 8.2e-111;
Matches 1067; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

QY 1 ATGACTGACGACATGTTGACGCTGATGAAATGGCACACGCGAGAAAGATTATTCGCCG 60
DB 1 ATGACTGACGACATGTTGACGCTGATGAAATGGCACATGCTGAGAAAGAAATATTCGCCG 60
QY 61 TTTTGGGATGCGGAGATGACCGCGCGCAAAAGAGCTTGGCGCTGATGGCAGAAGC 120
DB 61 TTTTGGGATGCGGAGATGACCGCGCGCAAAAGAGCTTGGCGCTGATGGCAGAAGC 120
QY 61 TTTTGGGATGCGGAGATGACCGCGCGCAAAAGAGCTTGGCGCTGATGGCAGAAGC 120
DB 61 TTTTGGGATGCGGAGATGACCGCGCGCAAAAGAGCTTGGCGCTGATGGCAGAAGC 120
QY 121 AATGTCGATGCGGCGCTGCTTCACTCTTATCACTGATCACTACTATTCGGCTGGCTG 180
DB 121 AATGTCGATGCGGCGCTGCTTCACTCTTATCACTGATCACTACTATTCGGCTGGCTG 180
QY 181 TACTGCTATTTGCGACGCAAGTACGCGATGTCATGACACCAACGCGACAGCAT 240
DB 181 TACTGCTATTTGCGACGCAAGTACGCGATGTCATGACACCAACGCGACAGCAT 240
QY 241 TGGGCGGCGATGACCGCGCGCAAGCTTGGCGCTGATGGCAGAAGC 300
DB 241 TGGGCGGCGATGACCGCGCGCAAGCTTGGCGCTGATGGCAGAAGC 300
QY 301 ACCGATGCGGCGCGCAATTTCTATCGGCGCGTGGCGAGCTGACACGCGCGCGCAAG 360
DB 301 ACCGATGCGGCGCGCAATTTCTATCGGCGCGTGGCGAGCTGACACGCGCGCGCAAG 360
QY 361 CGCATGCGGATGCGATGCGACAGCTCAATCTGCACTTCGCGCGCGAGCTCGAGAGCC 420
DB 361 CGCATGCGGATGCGATGCGACAGCTCAATCTGCACTTCGCGCGCGAGCTCGAGAGCC 420
QY 421 CTACCGGCGGCTGATGCTGCAATCAAGCCCTCGATGATGATGCGGACCATCAAG 480
DB 421 CTACCGGCGGCTGATGCTGCAATCAAGCCCTCGATGATGATGCGGACCATCAAG 480
QY 481 TGGCTCGAAGACAGAGTGAATCGCGAAGGCGCGCGTGTGTATCGTCCGCGCGCGCG 540
DB 481 TGGCTCGAAGACAGAGTGAATCGCGAAGGCGCGCGTGTGTATCGTCCGCGCGCGCG 540
QY 541 GCGTGGCGGCTGCGATCAAGCGCGCGTGGCGCGACATGAAGTGGCGATCGCACACC 600
DB 541 GCGTGGCGGCTGCGATCAAGCGCGCGTGGCGCGACATGAAGTGGCGATCGCACACC 600
QY 601 AATGCGATGATCGCGAGATCGCGCAATGCTTCCCTTCTGATGAGCTGATGACACTCG 660
DB 601 AATGCGATGATCGCGAGATCGCGCAATGCTTCCCTTCTGATGAGCTGATGACACTCG 660
QY 661 ACTGCTTCAATCGCGGATCAACCGAGCGCGCGCAATCGCTCAACCAACCGCATC 720

DB 661 ACTGCTTCAATCGCGGATCAACCGAGCGCGCGCAATCGCTCAACCAACCGCATC 720
QY 721 GTGCAATCGCGGACATCTCTTGGCTCAACACTTCCCGATGATCTTGGCTCTACACC 780
DB 721 GTGCAATCGCGGACATCTCTTGGCTCAACACTTCCCGATGATCTTGGCTCTACACC 780
QY 781 GCGCTGAGGACAGCGCTTCTGCGACATGTCGATGACGCGCGCTCGACATCTGAGAG 840
DB 781 GCGCTGAGGACAGCGCTTCTGCGACATGTCGATGACGCGCGCTCGACATCTGAGAG 840
QY 841 AAGAACGTGCGCGTGTGATGCGCGCGCTCGAGCTGATCAAGCGCGCGCGCTGCAAG 900
DB 841 AAGAACGTGCGCGTGTGATGCGCGCGCTCGAGCTGATCAAGCGCGCGCGCTGCAAG 900
QY 901 GACATGCGCATGAGCTCAACGAGATGATCCGCGAGTGGGACCTGCTGAAGTACCGCTCC 960
DB 901 GATATGCGCGCTGCAATCAACGAGATGATCCGCGAGTGGGATCTGCTGAAGTACCGCTCC 960
QY 961 TTCGCTATGCGCATCTTTCGCGCGTGTGCTGCTGCTCACTACAGCTGCGAGCGCGCGT 1020
DB 961 TTCGCTATGCGCATCTTTCGCGCGTGTGCTGCTGCTCACTACAGCTGCGAGCGCGCGT 1020
QY 1021 GAGCTGCGGACGACATGCAACCGAGCTGAAGCCCGGATGCTTCCATGAGCGG 1080
DB 1021 GAGCTGCGGACGACATGCAACCGAGCTGAAGCCCGGATGCTTCCATGAGCGG 1080
QY 1081 ATGCTGATGCTGCGGAGGATGCGCGGCTGCGCGCTATGCGACGACGATCTCG 1140
DB 1081 ATGCTGATGCTGCGGAGGATGCGCGGCTGCGCGCTATGCGACGACGATCTCG 1140
QY 1141 ATGCTGAGGAGGAGGATGCGGAGGATGCGGAGGATGCGGAGGATGCGGAGGATGCG 1200
DB 1141 ATGCTGAGGAGGAGGATGCGGAGGATGCGGAGGATGCGGAGGATGCGGAGGATGCG 1200
QY 1201 ATCATCCGCAAC 1212
DB 1201 ATCATCCGCAAC 1212

RESULT 15
AX721953 1212 bp DNA linear PAT 07-MAY-2003
LOCUS Sequence 21 from Patent EP1298213.
DEFINITION
ACCESSION AX721953
VERSION AX721953.1 GI:30422529
KEYWORDS
SOURCE
ORGANISM
synthetic construct
other sequences; artificial sequences.

REFERENCE
1 Shao, Z., Schumack, R., Kratzsch, P., Kenkles, J. and Weisser, H.
AUTHORS
TITLE
VARIANTS OF AN ERWINIA-TYPE CREATINASE
JOURNAL
Roche Diagnostics GmbH (DE) ; F.HOFFMANN-LA ROCHE AG (CH)
Location/Qualifiers

1..1212
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="variant CT2d7"
1..1212
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAD90220.1"
/db_xref="GI:30422524"
/translation="MTDMLHVMKHNKEKYSPESDAEMTRPSDVRMAENDVDA
ALFTSYHCINYSGLICYTFRKRYGAVITDQHAATTISAGIDGQPRRSFGNITTTD
WRNDYQAVRQLTPGARRIIGLEFHDVLDPEFRTLEBALPGVEFVIGPSPMWRIVK
SLEBQCLIREGARI CDVGGA CVAVKAGVPEHEVAIATNVAVEIASKSPFVELMD
TWTWFSQINTDGAHPVTRNI VOSGDILSLNTPMI PGYTALERTLCFDHVDASL
DTWEKNVAVHRRLGELIKPGARKDIALELNEMRYEMDLKXRSFGYGSFGVLSHY
GRBAGVELREDIDTVLQPGMVSMEMWMLPEGA PGAGYREHDLIIVGDEAGNITG
PFGPEPHNITRN"

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 7, 2005, 19:27:19 ; Search time 728 Seconds
(without alignments)
9855.394 Million cell updates/sec

Title: US-10-807-228a-2

Perfect score: 1212

Sequence: 1 ATGACTGACGACATGTTGCA.....AACACAACATCATCCGAC 1212

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

N_Geneseq_16Dec04:*
1: geneseqn19808:*
2: geneseqn19908:*
3: geneseqn20008:*
4: geneseqn20018:*
5: geneseqn20028:*
6: geneseqn20038:*
7: geneseqn20048:*
8: geneseqn20058:*
9: geneseqn20068:*
10: geneseqn20078:*
11: geneseqn20088:*
12: geneseqn20098:*
13: geneseqn20108:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Query Match	Length	DB	ID	Description
1	1210.4	99.9	1212	2	AAT61367	Aat61367 Creatinin
2	1210.4	99.9	1212	2	AAT38807	Aat38807 Creatinin
3	1210.4	99.9	1212	2	AAV5569	Aav5569 Strable cr
4	1210.4	99.9	1212	4	AA514742	Aa514742 Creatine
5	1210.4	99.9	1212	6	AA199856	Aa199856 Alkaligen
6	1210.4	99.9	1212	6	ABA93696	Ab93696 Creatine
7	1188	98.0	1215	2	AAT13291	Aat13291 Creatine
8	984.8	81.3	1215	8	ACC69514	Acc69514 Erwinda s
9	984.8	81.3	1215	8	ACC69519	Acc69519 Mutant Er
10	981.6	81.0	1215	8	ACC69515	Acc69515 Mutant Er
11	980	80.9	1212	8	ACC69517	Acc69517 Mutant Er
12	980	80.9	1212	8	ACC69516	Acc69516 Mutant Er
13	980	80.9	1212	8	ACC69518	Acc69518 Mutant Er
14	978.4	80.7	1212	8	ACC69521	Acc69521 Mutant Er
15	976.8	80.6	1212	8	ACC69520	Acc69520 Mutant Er
16	643.4	53.1	1209	2	AAT05502	Aat05502 Thermosta
17	638.8	52.7	1212	1	AAAB0274	AaAB0274 Sequence
18	637.2	52.6	1212	1	AAAB1271	AaAB1271 Sequence
19	564.6	46.6	1134	1	AAAT0877	AaAT0877 Sequence
20	459.4	37.9	1233	1	AAAN90635	AaAN90635 Sequence

ALIGNMENTS

21	319.2	26.3	1282	2	AAV66090	AAV66090 Arthrobac
22	74.8	6.2	858	11	ABD05567	ABD05567 Pseudomon
23	74.8	6.2	963	11	ABD05831	ABD05831 Pseudomon
24	74.8	6.2	1176	11	ABD05710	ABD05710 Pseudomon
25	73.2	6.0	786	10	ADG73350	ADG73350 Pseudomon
26	73.2	6.0	786	10	ADG73352	ADG73352 Pseudomon
27	68	5.6	985	6	ABQ44816	ABQ44816 Oligonuc
28	68	5.6	985	6	ABQ44817	ABQ44817 Oligonuc
29	68	5.6	3157	2	AAV08896	AAV08896 Cellulose
30	67.8	5.6	1227	6	ABE1294	ABE1294 N. unifor
31	67.2	5.5	65140	4	AAAD17184	AAAD17184 Streptomy
32	67.2	5.5	125401	4	AAAD17186	AAAD17186 Streptomy
33	66.6	5.5	2052	8	ACA23861	ACA23861 Prokaryot
34	65.2	5.4	690	3	AAA05948	AAA05948 Humanised
35	65.2	5.4	2712	8	ACA36793	ACA36793 Prokaryot
36	65.2	5.4	110000	6	ABR03041_18	ABR03041_18 o
37	65	5.4	2253	13	ADG91713	ADG91713 Polyketid
38	64.6	5.3	164051	13	ADG91695	ADG91695 Polyketid
39	64.6	5.3	9546	12	ADN40876	ADN40876 plasmid p
40	64.4	5.3	2037	8	ACC44573	ACC44573 Glucoamyl
41	64	5.3	1284	2	AAZ06824	AAZ06824 Streptomy
42	64	5.3	1284	4	AAH74537	AAH74537 Nucleotid
43	63.6	5.2	690	6	AAH46285	AAH46285 Aemonia
44	63.4	5.2	1161	8	ACA23536	ACA23536 Prokaryot
45	63.2	5.2	1284	8	ADA70934	ADA70934 Rice gene

RESULT 1

ID AAT61367 standard; DNA; 1212 BP.

XX AAT61367;

XX 17-APR-1997 (first entry)

XX Creatinine amidinohydrolyase coding sequence.

XX Thermal stability; creatinine amidinohydrolyase; creatine; sarcosine;

XX urea; blood, ds.

XX Alcaligenes faecalis.

XX JP08308579-A.

XX 26-NOV-1996.

XX 16-MAY-1995; 95UP-00117283.

XX 16-MAY-1995; 95UP-00117283.

XX (TOYM) TOYOB KK.

XX WPI, 1997-059698/06.

XX P-PSDB; AAW11861.

XX Gene coding for creatinine amidinohydrolyase - used to quantify blood or

XX urinary creatinine as a disease indicator.

XX Claim 4; Page 10-11; 12pp; Japanese.

XX This sequence encodes a thermally stable creatinine amidinohydrolyase

XX which has a low Km value for creatine. The creatinine amidinohydrolyase

XX has the following physicochemical properties: (a) action: converts

XX creatine and water into sarcosine and urea; (b) optimum temp.: 40-45

XX deg.C; (c) optimum pH: 8.0-9.0; (d) thermal stability: about 50 or 1less

XX (pH 7.5, 30 mins); (e) stable pH: a Ph range of about 4-10; (f) Km value:

XX about 15.2 mM; (g) mol. wt.: about 4300 (SDS-PAGE); and (h) isoelectric

XX point: about 3.5. The creatinine amidinohydrolyase may be used in the

XX quantification of blood or urinary creatinine and creatine as an

XX indicator of various diseases

```

XX Sequence 1212 BP, 244 A; 400 C; 358 G; 210 T; 0 U; 0 Other;
SQ Query Match 99.9%; Score 1210.4; DB 2; Length 1212;
Best Local Similarity 99.9%; Pred. No. 4,2e-196;
Matches 1211; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGACTGACGATGTTGACGTTGATGAAATGACCAACGAGGAGAAATTAATTCGCGG 60
DB 1 ATGACTGACGATGTTGACGTTGATGAAATGACCAACGAGGAGAAATTAATTCGCGG 60
QY 61 TTTTGGATGCGGAGATGACCGCGCGCAAAAGACGTTGCGGCTGATGAGCAAGAAC 120
DB 61 TTTTGGATGCGGAGATGACCGCGCGCAAAAGACGTTGCGGCTGATGAGCAAGAAC 120
QY 121 AATGTCGATGCGGCGCTGTTCACTCTTATCACTGATCAACTACTATTCCGGCTGGCTG 180
DB 121 AATGTCGATGCGGCGCTGTTCACTCTTATCACTGATCAACTACTATTCCGGCTGGCTG 180
QY 181 TACCTGCTATTTGGACGCAAGTACGAGTATGATGACCAACGACGCAAGCAAT 240
DB 181 TACCTGCTATTTGGACGCAAGTACGAGTATGATGACCAACGACGCAAGCAAT 240
QY 241 TCGGCGGCGGATGACGCGCGCGGAGCCCTGCGCGCGGAGCTTCGCGGCAACATCACTAC 300
DB 241 TCGGCGGCGGATGACGCGCGCGGAGCCCTGCGCGCGGAGCTTCGCGGCAACATCACTAC 300
QY 301 ACCGACTGCGCGCGGCAATTTCTATGCGCGCGGCTGCGGCACTGACACGCGGCGCAAG 360
DB 301 ACCGACTGCGCGCGGCAATTTCTATGCGCGCGGCTGCGGCACTGACACGCGGCGCAAG 360
QY 361 CGCATGCGGATGAGTTCACCAACGATCAGTTCGCTTCGCGCGGCTGAGGAAAGCC 420
DB 361 CGCATGCGGATGAGTTCACCAACGATCAGTTCGCTTCGCGCGGCTGAGGAAAGCC 420
QY 421 CTACCGGCGGCTGAGTTCGTCGACATCAGCAGCCCTCGATGATGATGCGACATCAAG 480
DB 421 CTACCGGCGGCTGAGTTCGTCGACATCAGCAGCCCTCGATGATGATGCGACATCAAG 480
QY 481 TCGCTCGAAGAGCAGAAAGCTGATCCGCGAGGCGCGCGCTGCTGATGATGATGATGATG 540
DB 481 TCGCTCGAAGAGCAGAAAGCTGATCCGCGAGGCGCGCGCTGCTGATGATGATGATGATG 540
QY 541 GCGTCCGCGGCTGCGATCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
DB 541 GCGTCCGCGGCTGCGATCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
QY 601 AATGCGATGATCCGCGAGATCCGCAAAATGCTTCCCTTCTGAGAGCTGATGACACTGG 660
DB 601 AATGCGATGATCCGCGAGATCCGCAAAATGCTTCCCTTCTGAGAGCTGATGACACTGG 660
QY 661 ACTGCTTCAGTCCGCGATCAACACCGAGCGCGCGCAATCCGCTCACCAACCGCATC 720
DB 661 ACTGCTTCAGTCCGCGATCAACACCGAGCGCGCGCAATCCGCTCACCAACCGCATC 720
QY 721 GTGCAATCCGCGAGATCCCTTTCGCTCAACACCTTCCGATATCTTCGCTCACTACAC 780
DB 721 GTGCAATCCGCGAGATCCCTTTCGCTCAACACCTTCCGATATCTTCGCTCACTACAC 780
QY 781 GCGCTGAGACGACCTGTTCTGCGACATGTCGATGACGCGAGCTCGACATCTGGAG 840
DB 781 GCGCTGAGACGACCTGTTCTGCGACATGTCGATGACGCGAGCTCGACATCTGGAG 840
QY 841 AAGAAAGTGGCGGTGATGCGCGCGGCTGAGCTGATCAAGCCGCGCGCGCTGGCAG 900
DB 841 AAGAAAGTGGCGGTGATGCGCGCGGCTGAGCTGATCAAGCCGCGCGCGCTGGCAG 900
QY 901 GACATGCGCATGACCTCAACAGATGTAACCGGAGTGGACCTGCTGAAGTACGCTCC 960
DB 901 GACATGCGCATGACCTCAACAGATGTAACCGGAGTGGACCTGCTGAAGTACGCTCC 960
QY 961 TTCGCTATGCGCACTCTTCGCGGCTGCTGCTCACTACTACGCTGCGAGGCGCGGCTG 1020
DB 961 TTCGCTATGCGCACTCTTCGCGGCTGCTGCTCACTACTACGCTGCGAGGCGCGGCTG 1020

```

```

DB 961 TTCGCTATGCGCACTCTTCGCGGCTGCTGCTCACTACTACGCTGCGAGGCGCGGCTG 1020
QY 1021 GAGCTGCCCGAGACATGACACCGAGCTGAAGCCCGGCAATGCTGCTCCATGAGCCG 1080
DB 1021 GAGCTGCCCGAGACATGACACCGAGCTGAAGCCCGGCAATGCTGCTCCATGAGCCG 1080
QY 1081 ATGCTGATGCTGCGCGAGGAGATGCGCGGCTGCGCGGCTGATGCGAGCAGACATCTCG 1140
DB 1081 ATGCTGATGCTGCGCGAGGAGATGCGCGGCTGCGCGGCTGATGCGAGCAGACATCTCG 1140
QY 1141 ATGCTGCGGAGAGACGCTGCGAGAACATCAACCGGCTTCCTGCTGCTGCGAACAAC 1200
DB 1141 ATGCTGCGGAGAGACGCTGCGAGAACATCAACCGGCTTCCTGCTGCTGCGAACAAC 1200
QY 1201 ATCATCCGCAAC 1212
DB 1201 ATCATCCGCAAC 1212

RESULT 2
AAT38807
ID AAT38807 standard; DNA; 1212 BP.
XX
AC AAT38807;
XX
DT 17-OCT-2003 (revised)
DT 02-MAR-1998 (first entry)
XX
DB Creatine amidinohydrolyase gene.
XX
KM Creatine amidinohydrolyase enzyme; sarcosine; urea; dye; Km; ds.
OS Alcaligenes faecalis; - strain TE3581 (FERM P-14237).
XX
FH Key Location/Qualifiers
FT CDS 1..1212
FT /tag= a
FT /product= "Creatine amidinohydrolyase"
FT /transl_except= (pos:433..435,aa:Glu)
FT /note= "no stop codon given"
FT
XX
EP790303-A1.
XX
PD 20-AUG-1997.
XX
PF 13-FEB-1997; 97EP-00102270.
XX
PR 13-FEB-1996; 96JP-00025435.
XX
PA (TOYM ) TOYO BOSEKI KK.
XX
PI Sogabe A, Hattori T, Nishiya Y, Kawamura Y;
DR WPI; 1997-404731/38.
DR P-PSDB; AAM22893.
XX
PT Creatine amidinohydrolyase enzyme with low Km - for use in assay for
PT creatine.
XX
PS Disclosure; Page 14-15; 21pp; English.
XX
CC A novel creatine amidinohydrolyase enzyme has been developed which
CC catalyses the reaction of creatine with water to form sarcosine and urea,
CC is stable at temperatures of up to 50 degrees Celsius (pH 7.5, 30
CC minutes) and has an optimum temperature of 40-50 degrees Celsius, optimum
CC pH of 8-9, a Km value for creatine of 3.5-10 mM in a binding assay using
CC sarcosine oxidase and peroxidase, a molecular weight of 43 kD by SDS-PAGE
CC and an isoelectric point of 3.5. The present sequence encodes creatine
CC amidinohydrolyase derived from Alcaligenes faecalis strain TE3581 (FERM P-
CC 14237), which is the wild type creatine amidinohydrolyase to be mutated in
CC the present invention. The enzyme can be used to determine creatine in a
CC sample by measuring the absorbance of a dye formed by reacting the sample
CC with a reagent, comprising the enzyme, sarcosine oxidase and a

```